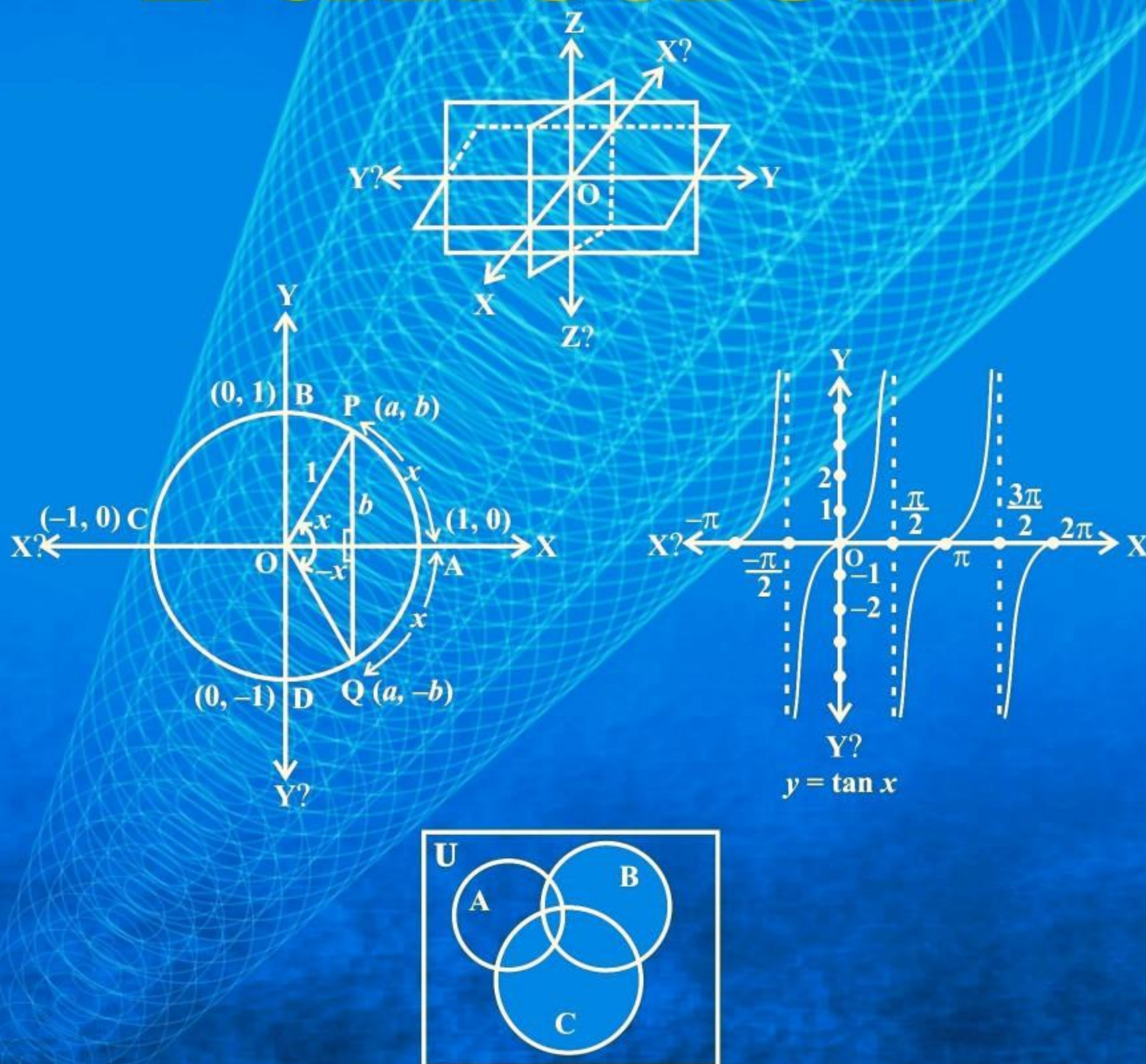
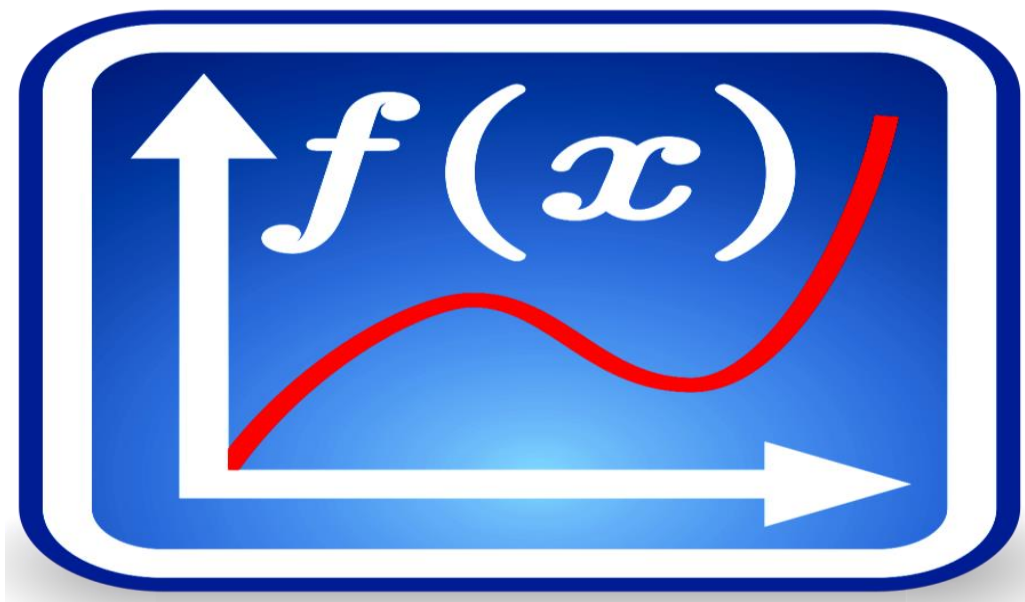
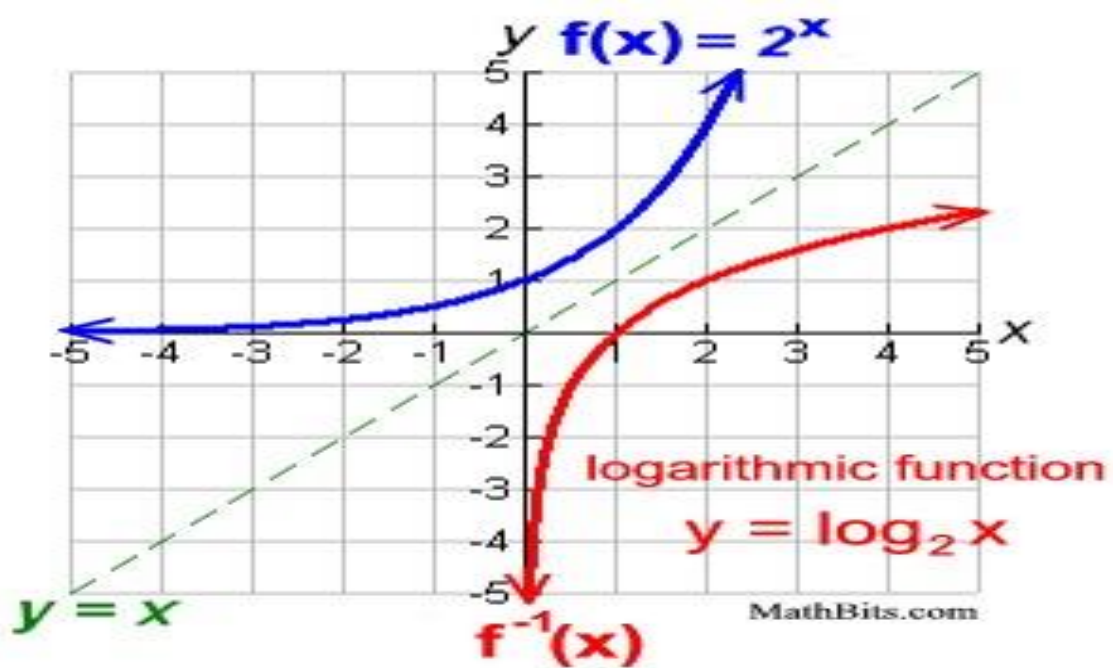


MATHEMATICS

Function



Nouh Hassan



All rights are reserved

Translated by:

Diab Publishing & Distribution



www.Diab-publishing.com

Preface

This book contains valuable lessons of scientific value in mathematics, especially in relation to the wide field of sports (functions).

This book is the unique encyclopedia of mathematics that helps in the teaching profession and in the culture of mathematics and in the contemporary aspects of this book with the various sciences of physics, engineering and technical sciences.

The researcher Mr. Nouh Hassan has relied on special curricula for educational education, research education and mathematics. It is a riches for the pioneers of mathematics and the world of the comprehensive scientific level.

Author Nouh Hassan

28/7/2019

Mr. Nouh Hassan in brief:



- Born in 1957, Maarake Village, Tyre in southern Lebanon.
- He received his primary and intermediate education at the official battle school and finished high school at the official secondary school.
- He graduated from the Teachers 'and Teachers' College in Nabatieh in 1979.
- He obtained his educational leave from the Lebanese University in mathematics in 1985.
- He was ranked first in Mathematics in the Lebanese Civil Service Council in 1995.
- He completed his graduation from the Faculty of Education at the Lebanese University in 1996 with the rank of secondary professor in the staff of the Ministry of National Education.
- Participated in many educational activities, training courses and public events.

Contents

| | |
|--|-----|
| 1. Introduction to functions..... | 1 |
| 2. Examples of function..... | 14 |
| 3. Functions and their graphs..... | 50 |
| 4. Lecture notes on relations and functions... | 125 |
| 5. Sets and function..... | 142 |
| 6. Numbers and functions..... | 193 |
| 7. Line and surface integrals functions..... | 313 |
| 8. Vector Algebra and functions..... | 359 |

Chapter one:

Introduction to functions

A function is a rule which operates on one number to give another number. However, not every rule describes a valid function. This unit explains how to see whether a given rule describes a valid function, and introduces some of the mathematical terms associated with functions.

In order to master the techniques explained here it is vital that you undertake plenty of practice exercises so that they become second nature.

After reading this text, and/or viewing the video tutorial on this topic, you should be able to:

- recognise when a rule describes a valid function,
- be able to plot the graph of a part of a function,
- find a suitable domain for a function, and find the corresponding range.

Contents

| | |
|-------------------------------------|---|
| 1. What is a function? | 2 |
| 2. Plotting the graph of a function | 3 |
| 3. When is a function valid? | 4 |
| 4. Some further examples | 6 |

1. What is a function?

Here is a definition of a function.

A function is a rule which maps a number to another unique number.

In other words, if we start off with an input, and we apply the function, we get an output.

For example, we might have a function that added 3 to any number. So if we apply this function to the number 2, we get the number 5. If we apply this function to the number 8, we get the number 11. If we apply this function to the number x , we get the number $x + 3$.

We can show this mathematically by writing

$$f(x) = x + 3.$$

The number x that we use for the input of the function is called the *argument* of the function. So if we choose an argument of 2, we get

$$f(2) = 2 + 3 = 5.$$

If we choose an argument of 8, we get

$$f(8) = 8 + 3 = 11.$$

If we choose an argument of -6 , we get

$$f(-6) = -6 + 3 = -3.$$

If we choose an argument of z , we get

$$f(z) = z + 3.$$

If we choose an argument of x^2 , we get

$$f(x^2) = x^2 + 3.$$

At first sight, it seems that we can pick any number we choose for the argument. However, that is not the case, as we shall see later. But because we do have some choice in the number we can pick, we call the argument the *independent variable*. The output of the function, e.g. $f(x)$, $f(5)$, etc. depends upon the argument, and so this is called the *dependent variable*.



Key Point

A function is a rule that maps a number to another unique number.

The input to the function is called the *independent variable*, and is also called the *argument* of the function. The output of the function is called the *dependent variable*.

2. Plotting the graph of a function

If we have a function given by a formula, we can try to plot its graph. Suppose, for example, that we have a function f defined by

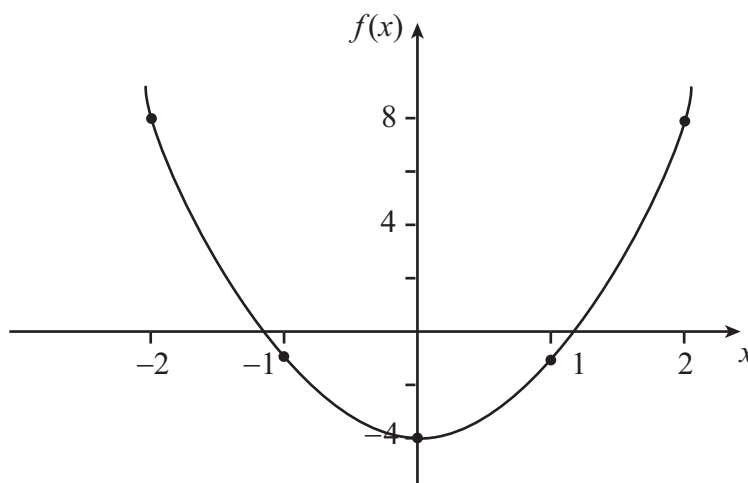
$$f(x) = 3x^2 - 4.$$

The argument of the function (the independent variable) is x , and the output (the dependent variable) is $3x^2 - 4$. So we can calculate the output of the function for different arguments:

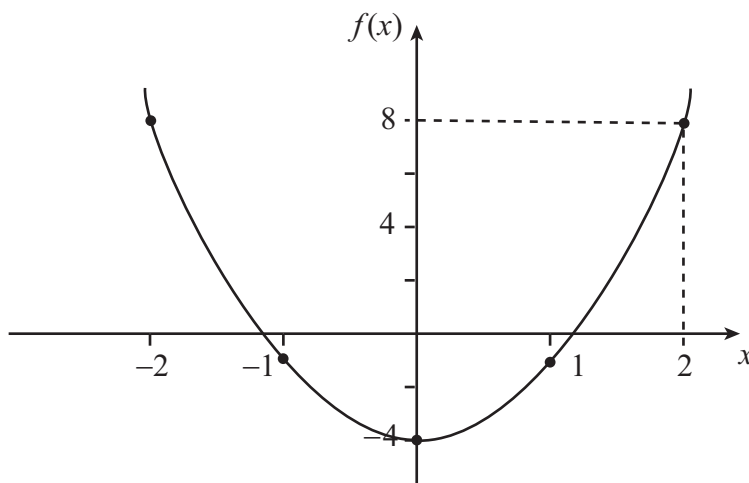
$$\begin{aligned} f(0) &= 3 \times 0^2 - 4 &= -4 \\ f(1) &= 3 \times 1^2 - 4 &= -1 \\ f(2) &= 3 \times 2^2 - 4 &= 8 \\ f(-1) &= 3 \times (-1)^2 - 4 &= -1 \\ f(-2) &= 3 \times (-2)^2 - 4 &= 8. \end{aligned}$$

We can put this information into a table to help us plot the graph of the function.

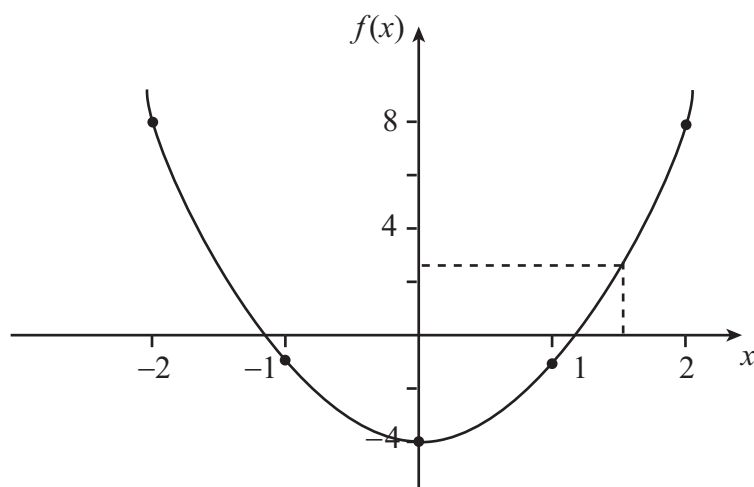
| x | $f(x)$ |
|-----|--------|
| -2 | 8 |
| -1 | -1 |
| 0 | -4 |
| 1 | -1 |
| 2 | 8 |



We can use the graph of the function to find the output corresponding to a given argument. For instance, if we have an argument of 2, we start on the horizontal axis at the point where $x = 2$, and we follow the line up until we reach the graph. Then we follow the line across so that we can read off the value of $f(x)$ on the vertical axis. In this case, the value of $f(x)$ is 8. Of course we already know this, because $x = 2$ is one of the values in our table.



But we can also use the graph for values of x which are not in our table. If we have an argument of 1.5, we follow the line up to the graph, and then across to the vertical axis. The result is a number between 2 and 3.



If we want to calculate this number exactly, we can substitute 1.5 into the formula:

$$f(1.5) = 2.75$$

3. When is a function valid?

Our definition of a function says that it is a rule mapping a number to another unique number. So we cannot have a function which gives two different outputs for the same argument. One easy way to check this is from the graph of the function, by using a ruler. If the ruler is aligned vertically, then it only ever crosses the graph once; no more and no less. This means that the graph represents a valid function.

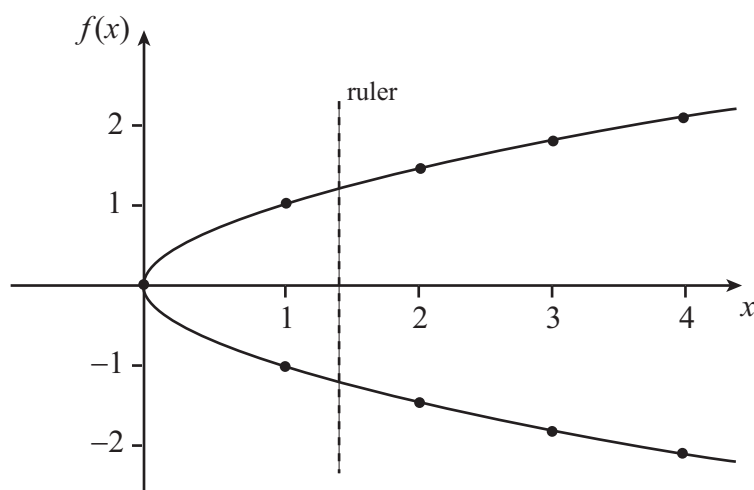
What happens if we try to define a function with more than one output for the same argument? Let's try an example. Suppose we try to define a function by saying that

$$f(x) = \sqrt{x}.$$

In the same way as before, we can produce a table of results to help us plot the graph of the function:

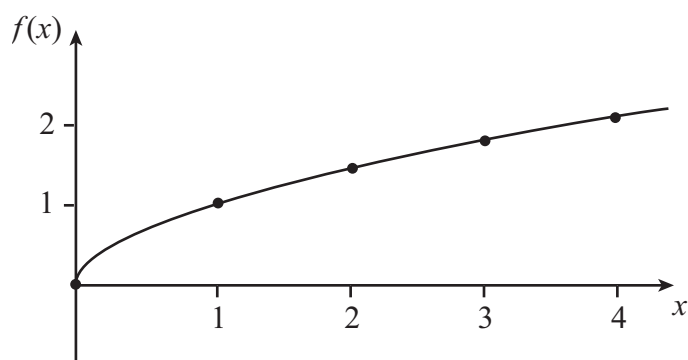
$$\begin{aligned} f(0) &= 0 \\ f(1) &= \pm 1 \\ f(2) &= \pm 1.4 && \text{to 1 d.p.} \\ f(3) &= \pm 1.7 && \text{to 1 d.p.} \\ f(4) &= \pm 2. \end{aligned}$$

(If we try to use any negative arguments, we end up in trouble because we are trying to find the square root of a negative number.) Plotting the results from the table, we get the following graph.



Using the ruler, it is quite clear that there are two values for all of the positive arguments. So as it stands, this is not a valid function.

One way around this problem is to define \sqrt{x} to take only the positive values, or zero: this is sometimes called the *positive square root* of x . However, there is still the issue that we cannot choose a negative argument. So we should also choose to restrict the choice of argument to positive values, or zero.



When considering these kinds of restrictions, it is important to use the right mathematical language. We say that the set of possible inputs is called the *domain* of the function, and the set of corresponding outputs is called the *range*. In the example above, we have defined the function as follows:

$$f(x) = \sqrt{x} \quad x \geq 0, \quad f(x) \geq 0,$$

so that the domain of the function is the set of numbers $x \geq 0$, and the range is the corresponding set of numbers $f(x) \geq 0$.



Key Point

The *domain* of a function is the set of possible inputs. The *range* of a function is the set of corresponding outputs.

4. Some further examples

Example

Consider the function

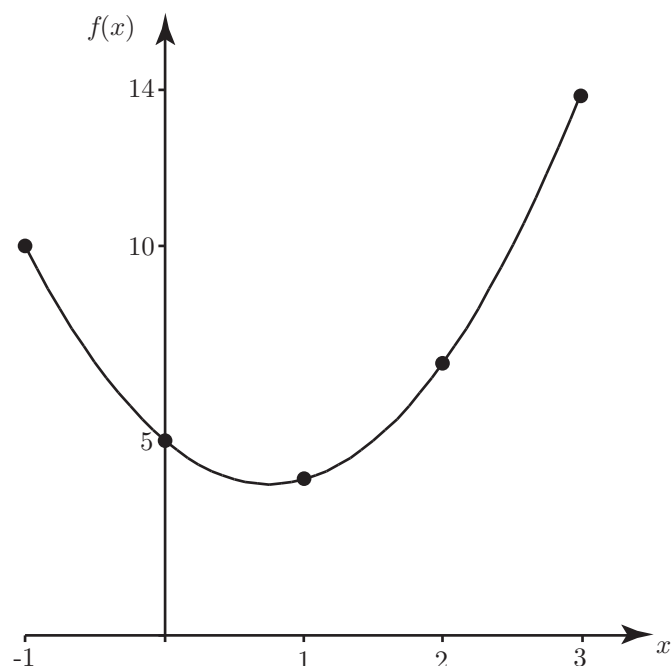
$$f(x) = 2x^2 - 3x + 5.$$

To make sure that the function is valid, we need to check whether we get exactly one output for each input, and whether there needs to be any restriction on the domain. As before, we can calculate the output of this function at some specific values to help us with plotting our graph:

$$\begin{aligned}f(0) &= 2 \times 0^2 - 3 \times 0 + 5 \\&= 5, \\f(1) &= 2 \times 1^2 - 3 \times 1 + 5 \\&= 2 - 3 + 5 \\&= 4, \\f(2) &= 2 \times 2^2 - 3 \times 2 + 5 \\&= 8 - 6 + 5 \\&= 7, \\f(3) &= 2 \times 3^2 - 3 \times 3 + 5 \\&= 18 - 9 + 5 \\&= 14, \\f(-1) &= 2 \times (-1)^2 - 3 \times (-1) + 5 \\&= 2 + 3 + 5 \\&= 10.\end{aligned}$$

Now we can put this into a table, and plot the graph.

| x | $f(x)$ |
|-----|--------|
| -1 | 10 |
| 0 | 5 |
| 1 | 4 |
| 2 | 7 |
| 3 | 14 |



A vertical ruler always crosses the graph once, and so the domain needs no restrictions, and the function is valid. We can also see from the graph that the minimum output occurs when $x = 0.75$, and that is when $f(x) = 3.875$. So the range of the function is $f(x) \geq 3.875$.

Example

What about the function

$$f(x) = \frac{1}{x}?$$

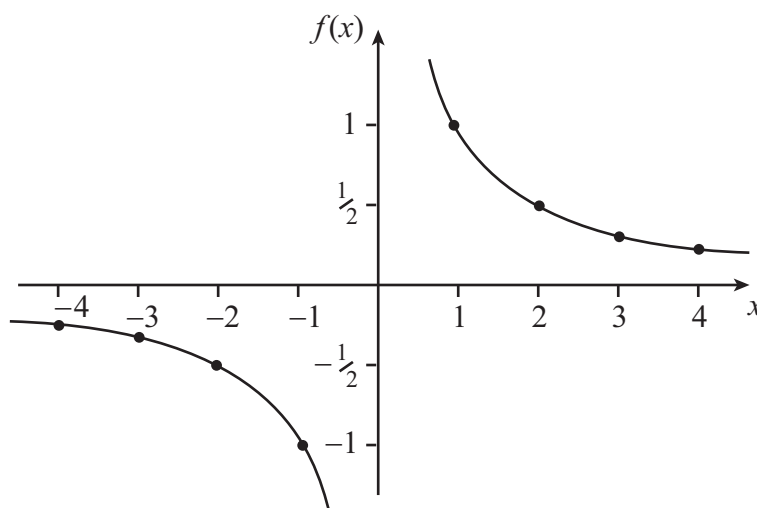
As usual, the first step is to check some values.

$$f(1) = 1, \quad f(2) = \frac{1}{2}, \quad f(3) = \frac{1}{3}, \quad f(4) = \frac{1}{4},$$

$$f(-1) = \frac{1}{(-1)} = -1, \quad f(-2) = \frac{1}{(-2)} = -\frac{1}{2},$$

$$f(-3) = \frac{1}{(-3)} = -\frac{1}{3}, \quad f(-4) = \frac{1}{(-4)} = -\frac{1}{4}.$$

When we try to calculate $f(0)$ we have a problem, because we cannot divide by zero. So we have to restrict the domain to exclude $x = 0$.



Because of this problem when $x = 0$, we have to restrict the domain to make the function valid. You can also see from the graph that there is no value of x where $f(x) = 0$, so zero is also excluded from the range. The function is therefore defined by

$$f(x) = 1/x \quad x \neq 0, \quad f(x) \neq 0.$$

You might want to know what exactly is going on at this point when $x = 0$. One way to find out is to look at what is happening very close to zero. So let's try some positive values for the argument getting closer and closer to zero, in order to see what happens:

$$f(1) = 1, \quad f\left(\frac{1}{2}\right) = \frac{1}{1/2} = 2, \quad f\left(\frac{1}{10}\right) = \frac{1}{1/10} = 10,$$

$$f\left(\frac{1}{1,000}\right) = \frac{1}{1/1,000} = 1,000, \quad f\left(\frac{1}{1,000,000}\right) = \frac{1}{1/1,000,000} = 1,000,000.$$

So you can see that as we approach zero from the right, the output approaches infinity.

Now let's try some negative values for the argument, getting closer and closer to zero from the left-hand side, in order to see what happens:

$$f(-1) = \frac{1}{(-1)} = -1, \quad f\left(-\frac{1}{2}\right) = \frac{1}{(-1/2)} = -2, \quad f\left(-\frac{1}{10}\right) = \frac{1}{(-1/10)} = -10,$$

$$f\left(-\frac{1}{1,000}\right) = \frac{1}{(-1/1,000)} = -1,000, \quad f\left(-\frac{1}{1,000,000}\right) = \frac{1}{(-1/1,000,000)} = -1,000,000.$$

You can see that as we approach zero from the left, the output approaches negative infinity. So in this case, approaching zero from the left is very different from approaching it from the right.

Example

For our final example, take the function

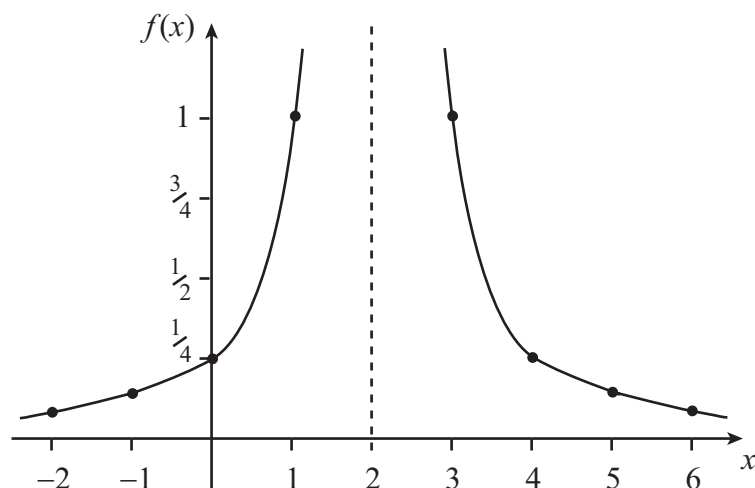
$$f(x) = \frac{1}{(x-2)^2}.$$

As usual, we must calculate some values:

$$\begin{aligned} f(-2) &= \frac{1}{(-2-2)^2} = \frac{1}{16} \\ f(-1) &= \frac{1}{(-1-2)^2} = \frac{1}{9} \\ f(0) &= \frac{1}{(0-2)^2} = \frac{1}{4} \\ f(1) &= \frac{1}{(1-2)^2} = 1 \\ f(2) &= \frac{1}{(2-2)^2} = \text{dividing by zero} \\ f(3) &= \frac{1}{(3-2)^2} = 1 \\ f(4) &= \frac{1}{(4-2)^2} = \frac{1}{4} \\ f(5) &= \frac{1}{(5-2)^2} = \frac{1}{9} \\ f(6) &= \frac{1}{(6-2)^2} = \frac{1}{16}. \end{aligned}$$

Now we can construct a table, and plot the graph of the function.

| x | $f(x)$ |
|-----|------------------|
| -2 | $\frac{1}{16}$ |
| -1 | $\frac{1}{9}$ |
| 0 | $\frac{1}{4}$ |
| 1 | 1 |
| 2 | dividing by zero |
| 3 | 1 |
| 4 | $\frac{1}{4}$ |
| 5 | $\frac{1}{9}$ |
| 6 | $\frac{1}{16}$ |



The vertical line at $x = 2$ is called an asymptote; it is a line which is approached by the graph but never reached. Although it is clear that we will have to omit $x = 2$ from the domain, this time the situation is slightly different. You can see that at either side of $x = 2$, $f(x)$ is approaching infinity. You should also notice that we cannot have any negative values for $f(x)$, or even $f(x) = 0$. So we have to define our function as

$$f(x) = \frac{1}{(x-2)^2} \quad x \neq 2, \quad f(x) > 0.$$

Exercises

1. Consider the function

$$f(x) = 2x^2 + 5x - 3.$$

- Write down the argument of this function.
- Write down the dependent variable in terms of the argument.
- Use a table of values to help you plot the graph of the function.
- From your graph, estimate $f(1.5)$.
- Use your function to calculate $f(1.5)$ exactly.
- Write down the domain and range of the function.
- Re-write the function with argument y .

2. Consider the function

$$f(x) = \frac{1}{(x-3)^2}.$$

- Plot the graph of the function.
- Write down the domain and range of the function.
- Re-write the function with argument z .
- Use your graph to estimate $f(1)$.
- Use the function to calculate $f(1)$ exactly.
- Write down another function where $x = 4$ has to be omitted from the domain.

3. Consider the function

$$f(x) = 3\sqrt{x}.$$

- What assumptions must be made about the function to ensure validity?
- Plot the graph of the function.

- (c) Write down the domain and range of the function.
 (d) Calculate $f(3)$.
 (e) What happens if you try to calculate $f(-2)$?

4. Consider the function

$$f(x) = \frac{1}{x}.$$

- (a) Plot the graph of the function.
 (b) Write down the domain and range of the function.
 (c) What happens to the output of the function as the argument approaches zero?
 (d) Is approaching zero from the left different to approaching zero from the right? If yes, why?
 (e) Calculate $f(2)$, $f(-10)$ and $f(z^3)$.

5. In the following list, you should write down the domain and range for each function, and then pair up functions that share the same domain and range.

$$f(x) = 2 \sin x - 1$$

$$f(x) = x^2 - 6x + 9$$

$$f(x) = 2e^{-x}$$

$$f(x) = 4 - x^2$$

$$f(x) = 2x - x^2 + 3$$

$$f(x) = 3e^{5x}$$

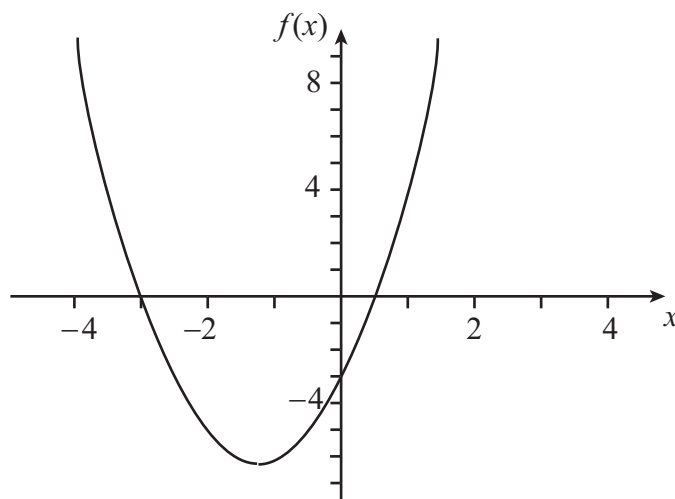
$$f(x) = 2 \cos 2x - 1$$

$$f(x) = 4x^2 - 16x + 16$$

Answers

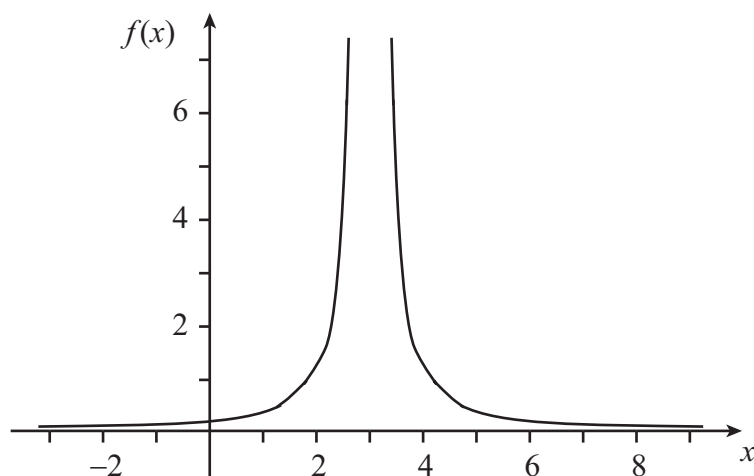
1.
 (a) The argument is x .
 (b) The dependent variable is $2x^2 + 5x - 3$ in terms of x .
 (c) An example of a table of values might be

| | | | | | | | |
|--------|----|----|----|----|----|---|----|
| x | -4 | -3 | -2 | -1 | 0 | 1 | 2 |
| $f(x)$ | 9 | 0 | -5 | -6 | -3 | 4 | 15 |



- (d) Draw a line up from $x = 1.5$, and read off the value on the y -axis.
- (e) $f(1.5) = 9$.
- (f) Reading off from the graph, $f(x) \geq -6.125$, and also $-\infty < x < \infty$.
- (g) The function is $f(y) = 2y^2 + 5y - 3$ using an argument y .

2.
(a)



(b) The domain is $x \neq 3$, the range is $f(x) > 0$.

(c) The function is $f(z) = \frac{1}{(z-3)^2}$ using an argument z .

(d) Draw a line up from $x = 1$, and read off the value on the y -axis.

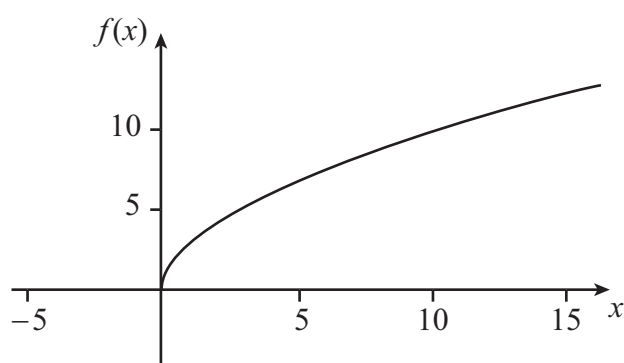
(e) $f(1) = \frac{1}{(1-3)^2} = \frac{1}{4}$.

(f) One such function is $f(x) = \frac{1}{(x-4)^2}$, but there are many others.

3.

(a) We must assume that we take the positive square root, and that we take $x \geq 0$.

(b)

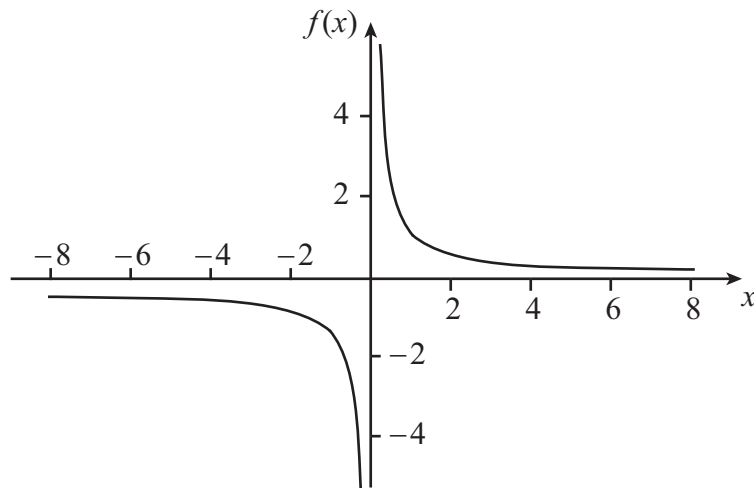


(c) The domain is $x \geq 0$, the range is $f(x) \geq 0$.

(d) $f(3) = \sqrt{3} = 1.732$ (to 1 d.p.).

(e) If you try to calculate $f(-2)$, you will be attempting to find the square root of a negative number. This has no real solutions.

4.
(a)



(b) The domain is $x \neq 0$, the range is $f(x) \neq 0$.

(c) The output approaches $\pm\infty$.

(d) Approaching from the left is different to approaching from the right. Approaching from the left, the output decreases rapidly to negative infinity. Approaching from the right, the output increases rapidly to positive infinity.

(e) $f(2) = \frac{1}{2}$, $f(-10) = -\frac{1}{10}$, $f(z^3) = 1/z^3$.

5.

| | domain | range |
|--------------------------|------------------------|-----------------------|
| $f(x) = 2 \sin x - 1$ | $-\infty < x < \infty$ | $-3 \leq f(x) \leq 1$ |
| $f(x) = x^2 - 6x + 9$ | $-\infty < x < \infty$ | $f(x) \geq 0$ |
| $f(x) = 2e^{-x}$ | $-\infty < x < \infty$ | $f(x) > 0$ |
| $f(x) = 4 - x^2$ | $-\infty < x < \infty$ | $f(x) \leq 4$ |
| $f(x) = 2x - x^2 + 3$ | $-\infty < x < \infty$ | $f(x) \leq 4$ |
| $f(x) = 3e^{5x}$ | $-\infty < x < \infty$ | $f(x) > 0$ |
| $f(x) = 2 \cos 2x - 1$ | $-\infty < x < \infty$ | $-3 \leq f(x) \leq 1$ |
| $f(x) = 4x^2 - 16x + 16$ | $-\infty < x < \infty$ | $f(x) \geq 0$ |

So the pairs are

$$f(x) = 2 \sin(x) - 1 \quad \text{and} \quad f(x) = 2 \cos 2x - 1,$$

$$f(x) = x^2 - 6x + 9 \quad \text{and} \quad f(x) = 4x^2 - 16x + 16,$$

$$f(x) = 2e^{-x} \quad \text{and} \quad f(x) = 3e^{5x},$$

$$f(x) = 4 - x^2 \quad \text{and} \quad f(x) = 2x - x^2 + 3.$$

Examples of Functions

In this document is provided examples of a variety of functions. The purpose is to convince the beginning student that functions are something quite different than polynomial equations.

These examples illustrate that the domain of functions need not be the real numbers or any other particular set. These examples illustrate that the rule for a function need not be an equation and may in fact be presented in a great variety of ways.

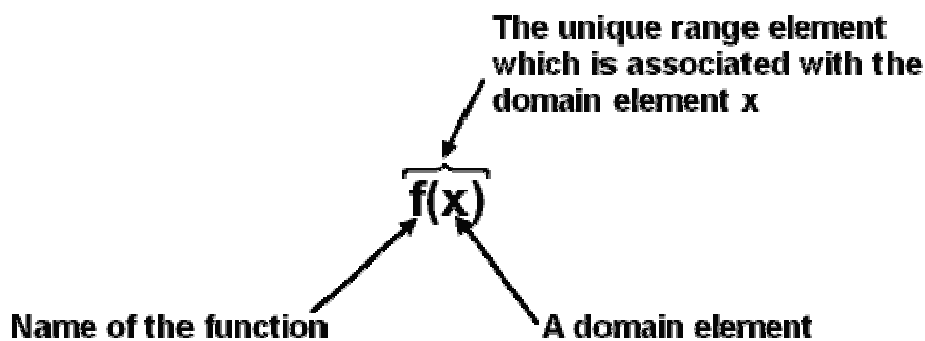
Some of the functions receive more complete treatment/discussion than others. If a particular function or function type is normally a part of a College Algebra course, then a more complete discussion of that function and its properties is likely.

Throughout this document the following definition of function, functional notation and convention regarding domains and ranges will be used.

Definition: A function consists of three things;

- i) A set called the domain
- ii) A set called the range
- iii) A rule which associates each element of the domain with a unique element of the range.

Functional Notation:



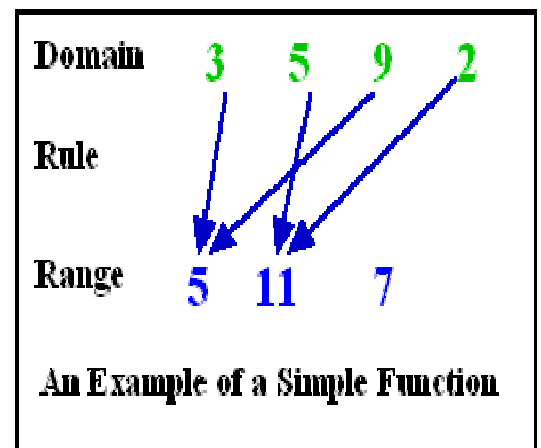
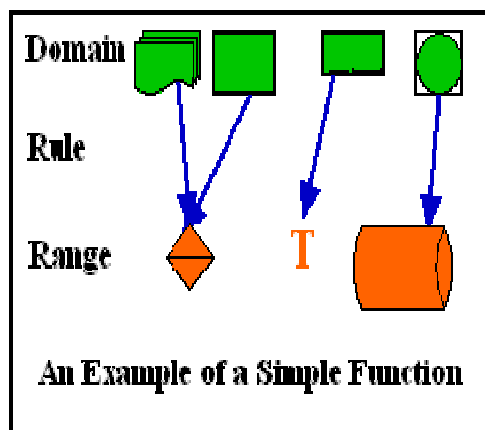
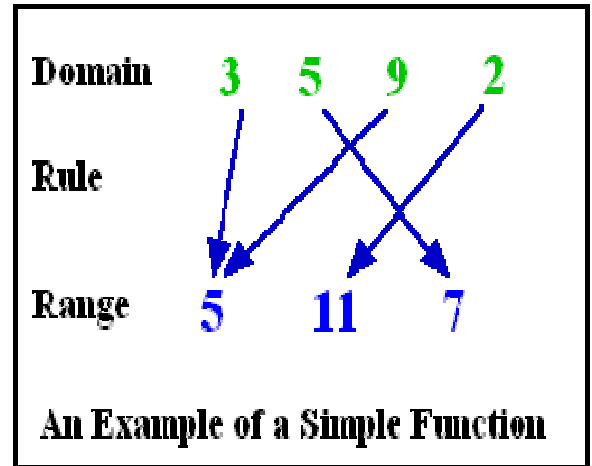
Convention: When the domain or range of a function is not specified, it is assumed that the range is \mathbf{R} , and the domain is the largest subset of \mathbf{R} for which the rule makes sense.

I) Some Very Simple Functions

The domains and rule of the three functions shown in these diagrams are quite simple, but they are indeed functions.

Moreover view of functions consisting of two sets with arrows from elements of the domain to elements of the range will serve well to help understand functions in general

The example below shows quite clearly that the domain and range of a function need not be sets of numbers. In most College Algebra courses the functions discussed do have domains and ranges which are sets of numbers, but that is not a requirement of the definition of function.



II) Linear Functions

Definition: A linear function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule can be expressed as a linear equation.

Recall a linear equation in two variables is an equation which can be written as $y = mx + b$ where m and b are real numbers.

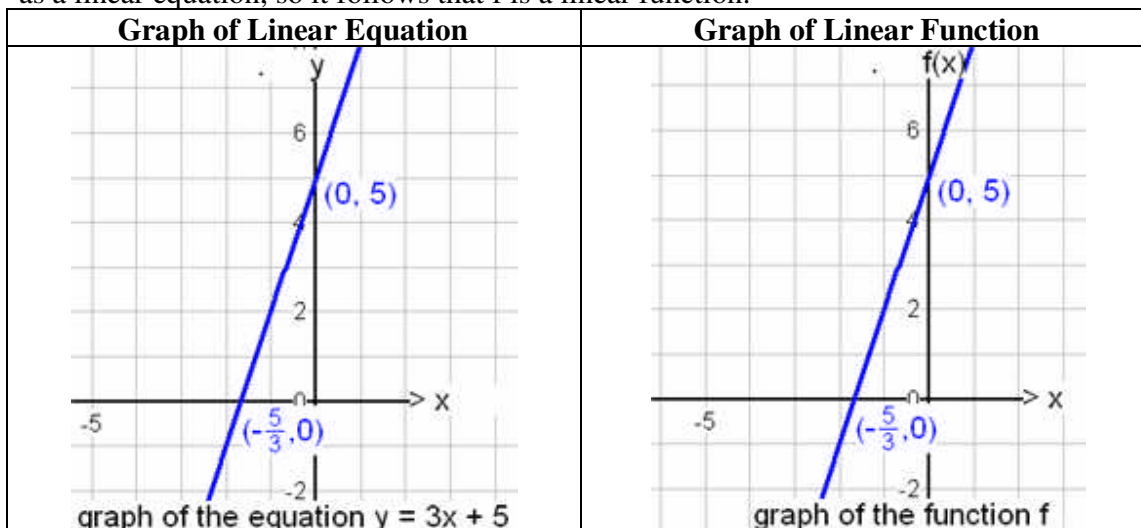
This leads to the following alternate **preferred** definition for a linear function.

Definition: A linear function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule can be written as $f(x) = mx + b$ where x is a domain element, $m \in \mathbf{R}$ and $b \in \mathbf{R}$.

Comment: Linear functions are a special case of the more general class of functions called polynomial functions. Thus the set of linear functions is a subset of the set of polynomial functions.

Example: Let f be the function whose rule is $f(x) = 3x + 5$.

Convention dictates that the range of this function is \mathbf{R} and because the rule makes sense for every real number substituted into the rule, the domain is also \mathbf{R} . The rule is written as a linear equation, so it follows that f is a linear function.



Comment: The word linear is now being used as an important adjective in four distinct contexts. We speak of linear polynomials, linear equations in one variable, linear equations in two variables, and linear functions.

- A linear polynomial is an expression which can be written as $ax + b$ where a and b are real numbers. A linear polynomial is a mathematical creature just like natural numbers, rational numbers, irrational numbers, matrices, and vectors are mathematical creatures. There is no equation involved. The concept of solving a linear polynomial is meaningless. The arithmetic-like concepts of addition, multiplication, and subtraction in the context of all polynomials as well as linear polynomials does have meaning.
- A linear equation in one variable is an equation which can be written as $ax + b = 0$ where a and b are real numbers. The graph of a linear equation in one variable is a single point on the x -axis unless $a = 0$ and $b \neq 0$ in which case there are no solutions and hence no graph.
- A linear equation in two variables is an equation which can be written in the form $y = mx + b$ where m and b are real numbers. This is called the slope-intercept form of the equation of a line. The graph of a linear equation in two variables is a line. The coefficient m is the slope of the graph and b is the y -intercept of that graph.
- A linear function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule may be written as $f(x) = mx + b$ where a and b are real numbers. The graph of a linear function is a non-vertical line with slope m and y -intercept b . A linear

function is a mathematical creature just like natural numbers, rational numbers, irrational numbers, matrices, vectors, and linear polynomials are mathematical creatures. The concept of solving a linear function is meaningless. The arithmetic-like concepts of addition, multiplication, and subtraction in the context of all functions as well as linear functions does have meaning. A variety of additional facts peculiar to functions will be studied in this course.

III) Quadratic Functions

Definition: A quadratic function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule can be expressed as a quadratic equation.

Recall that a quadratic equation in two variables is an equation that can be written $y = ax^2 + bx + c$ where a , b , and c are real numbers and $a \neq 0$.

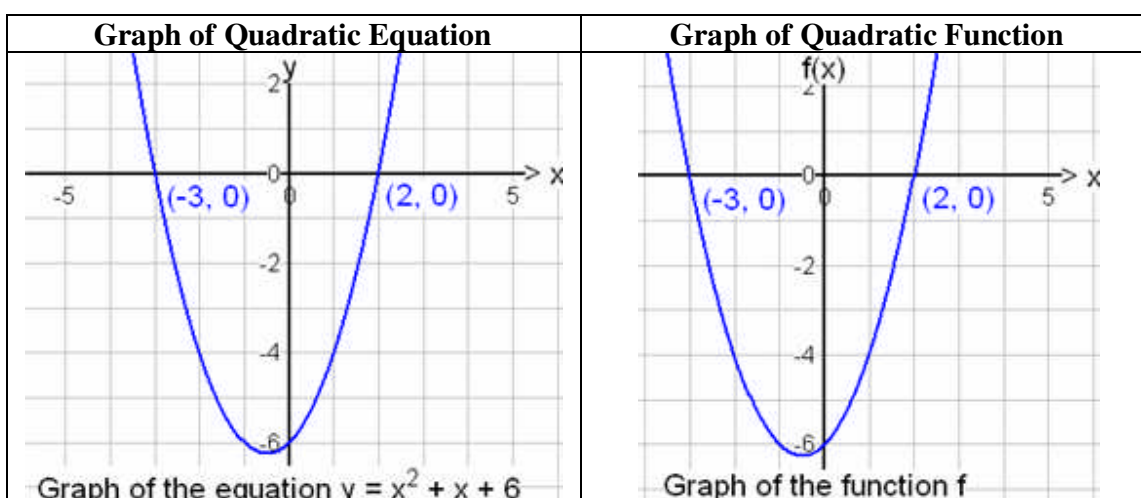
This leads to the following alternate **preferred** definition for a quadratic function.

Definition: A quadratic function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule can be written as $f(x) = ax^2 + bx + c$ where a , b , and c are real numbers and $a \neq 0$.

Comment: Quadratic functions are a special case of the more general class of functions called polynomial functions. Thus the set of quadratic functions is a subset of the set of polynomial functions.

Example: Let f be the function whose rule is $f(x) = x^2 + x - 6$.

The convention dictates that the range of this function is \mathbf{R} and because the rule makes sense for every real number substituted into the rule, the domain is also \mathbf{R} . The rule is written as a quadratic equation, so it follows that f is a quadratic function.



Comment: The word quadratic is now being used as an important adjective in four distinct contexts. We speak of quadratic polynomials, quadratic equations in one variable, quadratic equations in two variables, and quadratic functions.

- A quadratic polynomial is an expression which can be written as $ax^2 + bx + c$ where a and b are real numbers. A quadratic polynomial is a mathematical creature just like natural numbers, rational numbers, irrational numbers, matrices, vectors, linear polynomials, and linear functions are mathematical creatures. There is no equation involved. The concept of solving a quadratic polynomial is meaningless. The arithmetic-like concepts of addition, multiplication, and subtraction in the context of all polynomials as well as quadratic polynomials do have meaning.
- A quadratic equation in one variable is an equation which can be written as $ax^2 + bx + c = 0$ where a , b , and c are real numbers. The graph of a quadratic equation in one variable may be a single point, or two points on the x -axis. If the discriminant of the quadratic polynomial is negative the quadratic equation in one variable has no real solutions and hence has no graph.
- A quadratic equation in two variables is an equation which can be written in the form $y = ax^2 + bx + c$ where a , b , and c are real numbers and $a \neq 0$. The graph of a quadratic equation in two variables is a parabola which opens up if the leading coefficient is positive and opens down if the leading coefficient is negative.
- A quadratic function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule may be written as $f(x) = ax^2 + bx + c$ where a , b , and c are real numbers and $a \neq 0$. The graph of a quadratic function is a parabola which opens up if $a > 0$ and opens down if $a < 0$. A quadratic function is a mathematical creature just like natural numbers, rational numbers, irrational numbers, matrices, vectors, and linear polynomials are mathematical creatures. The concept of solving a linear function is meaningless. The arithmetic-like concepts of addition, multiplication, and subtraction in the context of all functions as well as quadratic functions do have meaning. A variety of additional facts peculiar to functions will be studied in this course.

IV) Polynomial Functions

Definition: A polynomial function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule can be expressed as a polynomial equation.

Recall that a polynomial equation in two variables is an equation that can be written $y = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ where n is a whole number and each of the coefficients a_i is a real number.

This leads to the following alternate **preferred** definition for a quadratic function.

Definition: A polynomial function is a function whose domain is \mathbf{R} , whose range is \mathbf{R} , and whose rule can be expressed as $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ where n is a whole number and each of the coefficients a_i is a real number.

Examples: Every linear function is a polynomial function.

Examples: Every quadratic function is a polynomial function.

Example: Let f be the function whose rule is $f(x) = x^5 + 3x^2 - 6x + 9$.

Convention dictates that the range of this function is \mathbf{R} and because the rule makes sense for every real number substituted into the rule, the domain is also \mathbf{R} . The rule is written as a polynomial equation, so it follows that f is a polynomial function.

Polynomial functions will be a major topic of study in College Algebra.

V) Absolute Value Function

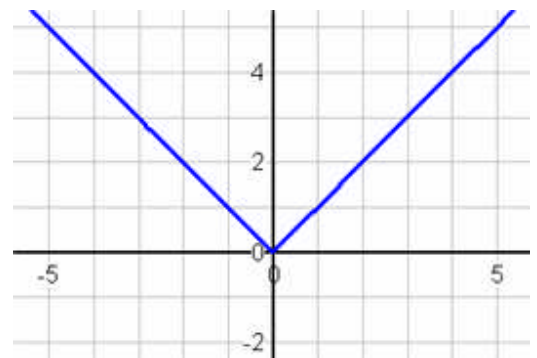
Definition: The absolute value function abs is the function whose rule is given by

$$abs(x) = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

Convention dictates that the range of this function is \mathbf{R} and because the rule makes sense for every real number substituted into the rule, the domain is also \mathbf{R} .

The graph of abs is shown at the right.

The name abs is less used in discussions involving the absolute value function. The more familiar notation $|x|$ is usually used in lieu of $abs(x)$.



VI) Sequences

Definition: A sequence is a function whose domain is the set of Natural Numbers \mathbf{N} .

The definition of sequence is pretty simply stated, but there are many special consequences of the fact that the domain of a sequence is \mathbf{N} . Sequences have been studied for centuries. Sequences have been studied with and without the concept of function. Quite a bit of special (mostly historical in origin) terminology and notation is used in a discussion of sequences.

When working with sequences, range elements are frequently called terms of the sequence. For example:

- The range element associated with the domain element 1 is called the first term of the sequence.

- The range element associated with the domain element 6 is called the sixth term of the sequence.
- The range element associated with the domain element n is called the n^{th} term of the sequence.

Because the domain of every sequence is \mathbf{N} , the graph of a sequence will consist of a set of discrete points.

Because the domain of every sequence is \mathbf{N} , it is possible to speak of the first domain element.

Because the domain of every sequence is \mathbf{N} , it is possible to speak of the next domain element.

Because the domain of every sequence is \mathbf{N} , it is not possible to speak of the last domain element.

Because the domain of every sequence is \mathbf{N} , it is possible to ask about and compute the sum of the first k terms. This is usually called the k^{th} partial sum of the sequence.

Definition: The n^{th} partial sum of a sequence is defined to be the sum of the first n terms of the sequence.

VI - A) Tau

The function whose name is the Greek letter τ (pronounced tau) is a function whose domain is the Natural Numbers \mathbf{N} . So τ is a sequence. The rule for τ is not given by a formula. The rule for τ is: $\tau(n)$ is the number of positive divisors of n .

To compute the range value associated with a particular domain element n , it is necessary to determine all positive divisors of n and simply count them. It is convenient to think of τ simply as a function which counts the number of positive divisors of domain elements.

$$\begin{array}{cccccc} \tau(1) = 1 & \tau(2) = 2 & \tau(3) = 2 & \tau(4) = 3 & \tau(5) = 2 & \tau(6) = 4 \\ \tau(7) = 2 & \tau(8) = 4 & \tau(9) = 3 & \tau(10) = 4 & \tau(11) = 2 & \tau(12) = 6 \end{array}$$

The graph of the first 12 terms of Tau consists of the points:

$$\begin{array}{cccccc} (1, 1) & (2, 2) & (3, 2) & (4, 3) & (5, 2) & (6, 4) \\ (7, 2) & (8, 4) & (9, 3) & (10, 4) & (11, 2) & (12, 6) \end{array}$$

Notice that if p is a prime number, then $\tau(p) = 2$ and if k is a composite number, then $\tau(k) > 2$. In fact sometimes τ is used to define prime numbers as:

Definition: A natural number p is prime if and only if $\tau(p) = 2$.

Notice how neatly this definition prohibits classifying 1 as a prime number.

VI - B) Sigma

The function whose name is the Greek letter σ (pronounced sigma) is a function whose domain is the Natural Numbers \mathbf{N} . So σ is a sequence. The rule for σ is not given by a formula. The rule for σ is: $\sigma(n)$ is the sum of the positive divisors of n . To compute the range value associated with a particular domain element n , it is necessary to determine all positive divisors of n and simply add them.

$$\begin{array}{llllll} \sigma(1) = 1 & \sigma(2) = 3 & \sigma(3) = 4 & \sigma(4) = 7 & \sigma(5) = 6 & \sigma(6) = 12 \\ \sigma(7) = 8 & \sigma(8) = 15 & \sigma(9) = 13 & \sigma(10) = 18 & \sigma(11) = 12 & \sigma(12) = 28 \end{array}$$

The graph of the first 12 terms of Sigma consists of the points:

$$\begin{array}{llllll} (1, 1) & (2, 3) & (3, 4) & (4, 7) & (5, 6) & (6, 12) \\ (7, 8) & (8, 15) & (9, 13) & (10, 18) & (11, 12) & (12, 28) \end{array}$$

VI - C) Fibonacci Sequence

Definition: The Fibonacci sequence F is the function whose domain is \mathbf{N} and whose rule is given recursively by: $F(1) = 1$, $F(2) = 1$, and for $n > 2$, $F(n) = F(n - 1) + F(n - 2)$

$$\begin{array}{llllll} F(1) = 1 & F(2) = 1 & F(3) = 2 & F(4) = 3 & F(5) = 5 & F(6) = 8 \\ F(7) = 13 & F(8) = 21 & F(9) = 34 & F(10) = 55 & F(11) = 89 & F(12) = 144 \end{array}$$

The graph of the first 12 terms of the Fibonacci sequence consists of the points:

$$\begin{array}{llllll} (1, 1) & (2, 1) & (3, 2) & (4, 3) & (5, 5) & (6, 8) \\ (7, 13) & (8, 21) & (9, 34) & (10, 55) & (11, 89) & (12, 144) \end{array}$$

VI - D) Arithmetic Sequences

Definition: An arithmetic sequence is a sequence whose consecutive terms have a common difference.

Equivalent Definition: An arithmetic sequence f is a function whose rule may be expressed as a linear equation of the form $f(n) = dn + b$ where d is the common difference and b is the difference $f(1) - d$.

Comments: Compare the equivalent definition of an arithmetic sequence with the definition of a linear function.

The domain of a linear function is \mathbf{R} The domain of an arithmetic sequence is \mathbf{N}

The rule for a linear function is $f(x) = mx + b$

The rule for an arithmetic sequence is $f(x) = dx + b$

The number b in the rule for an arithmetic sequence is the range value associated with 0 (if there were such a range element) so it corresponds exactly to the y-intercept of the linear function.

The common difference d is nothing more than the slope as you move from one range element to the next. The slope of the line joining two terms $(x, f(x))$ and $(x + 1, f(x + 1))$ of the sequence is given by $\frac{f(x+1) - f(x)}{(x+1) - x} = \frac{d}{1} = d$.

A casual approach is to view an arithmetic sequence as a linear function with domain \mathbf{N} .

Comment: Given any two pieces of information about an arithmetic sequence it is possible to determine its rule. The next three problem types illustrate the point.

Problem Type 1: If you are given the common difference and the first term of the arithmetic sequence, then it is possible to write the rule for the function. This is comparable to the slope-intercept situation/problem when working with linear functions.

Example: Suppose an arithmetic sequence named h has a common difference 8 and the first term is -5. Find the rule for the function h .

Solution: Since the function is an arithmetic sequence its rule is of the form $h(n) = dn + b$. In our case the common difference d is 8, the rule for h has the form $h(n) = 8n + b$. Because the first term is -5, $b = -5 - 8 = -13$ and the rule for the desired arithmetic sequence is given by $h(n) = 8n - 13$.

Problem Type 2: If you are given the common difference d and one term of an arithmetic sequence, then it is possible to write the rule for the function. This is comparable to the point-slope situation/problems when working with linear functions.

Example: Suppose an arithmetic sequence named h has a common difference 3 and the fifth term is 12. Find the rule for the function h .

Solution: Since the function is an arithmetic sequence its rule is of the form $h(n) = dn + b$. In our case the common difference d is 3, the rule for h has the form $h(n) = 3n + b$. Because the fifth term is 12, $h(5) = 12$, but according to the partially determined rule $h(5) = 3(5) + b = 15 + b$. These two representations for $h(5)$ yield the equation $12 = 15 + b$. Clearly then $b = -3$ and the rule for the desired arithmetic sequence is given by $h(n) = 3n - 3$.

Problem Type 3: If you are given two terms of an arithmetic sequence, then it is possible to write the rule for the function. This is comparable to the two point situation/problems when working with linear functions.

Example: Suppose the fourth term an arithmetic sequence named k is 10 and the seventh term is 28. Find the rule for the function h .

Solution: Since the function is an arithmetic sequence its rule is of the form $h(n) = dn + b$. The difference between the seventh and fourth terms is $3d$ and is also equal to $28 - 10 = 18$. That means $3d = 18$ and so the common difference d is 6. The rule for h has the form $h(n) = 6n + b$. Because the fourth term is 10, $h(4) = 10$, but according to the partially determined rule $h(4) = 6(4) + b = 24 + b$. These two representations for $h(4)$ yield the equation $10 = 24 + b$. Clearly then $b = -14$ and the rule for the desired arithmetic sequence is given by $h(n) = 6n - 14$.

Comment: The formula for the n^{th} partial sum of an arithmetic sequence named a is:

$$S_n = \frac{n}{2}(a_1 + a_n)$$

VI - D) Mod Functions

The function named mod_6

The name of this function is mod_6 . The domain of mod_6 is the set of all natural numbers. The range of mod_6 is the set $\{0,1,2,3,4,5\}$. Because the domain of mod_6 is \mathbf{N} , mod_6 is a sequence. The rule for mod_6 is given by:

$\text{mod}_6(n)$ is the remainder when n is divided by 6.

To see that mod_6 is a function, we refer back to the “arrow” concept of function. The fact that every natural number may be divided by 6 insures that an arrow emanates from each element of the domain.

The division algorithm states that for any natural number n there is a unique quotient q and a **unique** remainder r such that $r \in \{0,1,2,3,4,5\}$.

The fact that $r \in \{0,1,2,3,4,5\}$ insures that the arrows end in the range and the fact that the remainder is unique insures that only one arrow emanates for each domain element.

Therefore mod_6 is a function.

Here are examples of range elements associated with some domain elements

$$\text{mod}_6(3) = 3 \qquad \text{mod}_6(8) = 2 \qquad \text{mod}_6(17) = 5 \qquad \text{mod}_6(424) = 4$$

The function named mod_{11}

The name of this function is mod_{11} . The domain of mod_{11} is the set of all natural numbers. The range of mod_{11} is the set $\{0,1,2,3,4,5,6,7,8,9,10\}$. Because the domain of mod_{11} is \mathbf{N} , mod_{11} is a sequence. The rule for mod_{11} is given by:

$\text{mod}_{11}(n)$ is the remainder when n is divided by 11.

To see that mod_{11} is a function, we refer back to the “arrow” concept of function. The fact that every natural number may be divided by 11 insures that an arrow emanates from each element of the domain.

The division algorithm states that for any natural number n there is a unique quotient q and a **unique** remainder r such that $r \in \{0,1,2,3,4,5,6,7,8,9,10\}$.

The fact that $r \in \{0,1,2,3,4,5,6,7,8,9,10\}$ insures that the arrows end in the range and the fact that the remainder is unique insures that only one arrow emanates from each domain element.

Therefore mod_{11} is a function.

Here are examples of range elements associated with some domain elements

$$\text{mod}_{11}(3) = 3 \qquad \text{mod}_{11}(8) = 8 \qquad \text{mod}_{11}(17) = 6 \qquad \text{mod}_{11}(424) = 6$$

It should be clear that for each natural number k there is a corresponding mod_k function defined in the same manner as mod_6 and mod_{11} above.

VII) Exponential Function exp

e is one of those special numbers in mathematics, like pi, that keeps showing up in all kinds of important places. Like pi, e is an irrational number. The value of e may be approximated by $e \approx 2.7182818284$

This irrational number is the foundation of a very important pair of functions in mathematics. These two functions are exp and ln (that is the letter ell). The function exp is quite easy to define.

The function exp has domain \mathbf{R} and range \mathbf{R} . The rule for exp is given by the exponential equation $\exp(x) = e^x$. The graph of exp is shown here.

The function exp is called the exponential function base e .

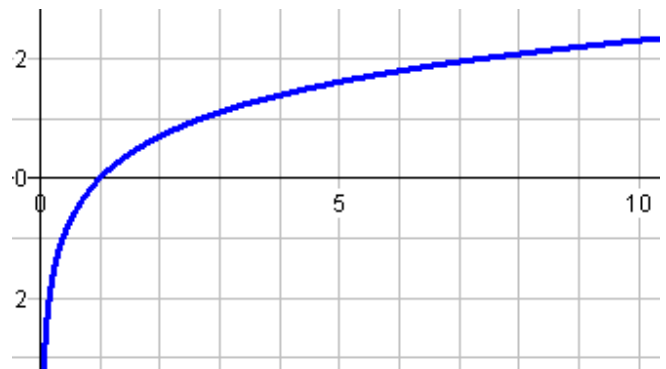
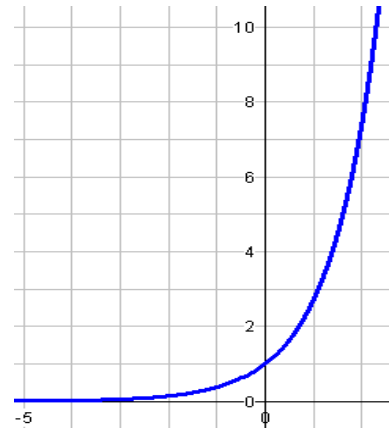
The function ln is called the natural logarithm function.

The domain of ln is all positive real numbers.

The rule for ln is given in terms of its inter-relation with the function exp. The function ln is the one and only function which has the property that

$$\ln(\exp(x)) = x \text{ and } \exp(\ln(x)) = x$$

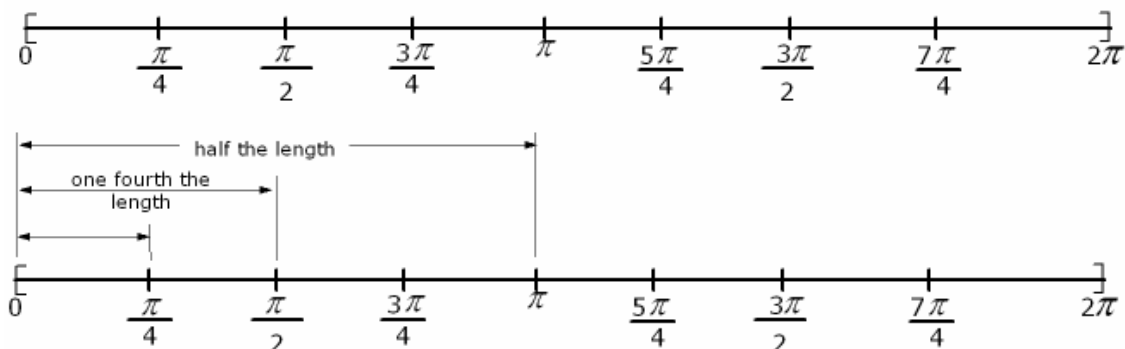
The graph of ln is shown at the right.



VIII) Circular Functions

We will now consider the functions whose names are sn, and cs, whose domain is the closed interval $[0, 2\pi]$ and whose range is the closed interval $[-1, 1]$. These and four other functions are called circular functions or more traditionally Trigonometric functions.

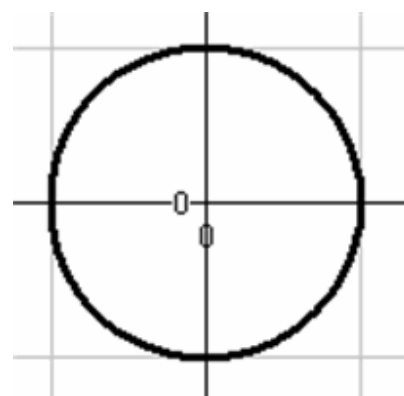
Observe the length of this domain is 2π . Here is a picture of the domain of the function named sn. A few points which we will use later in the discussion are marked.



The rule for this function will not be described with an equation but will instead be described in terms of the coordinates of points on the unit circle.

Recall that the unit circle is the circle with radius 1 whose center is at the origin of the Cartesian coordinate system and is described by the equation $x^2 + y^2 = 1$.

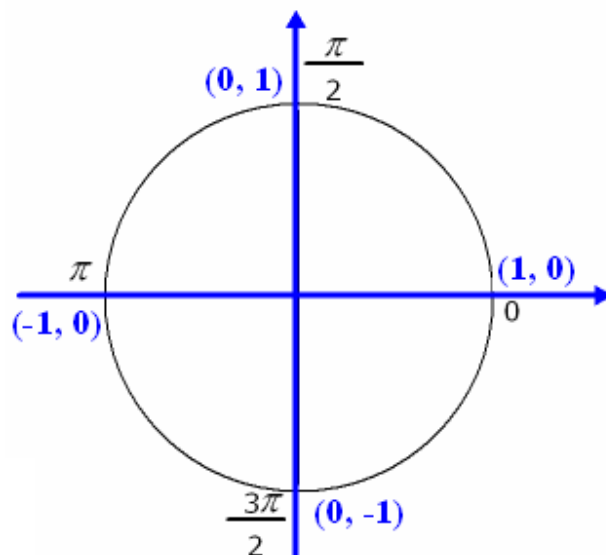
Also recall that the radius of a circle is given by the formula $C = 2\pi r$. In the case of the unit circle, the circumference is 2π . This circumference is exactly the same length as the domain of the function named sn.



The significance of this comparison is that for any real number in the domain of the function named sn, there is a corresponding point on the unit circle. The converse is also true, for every point on the circumference of the unit circle there is a real number in the domain of the function.

On the unit circle the point $(1, 0)$ is always considered the starting point and distance is always measured on the circumference in the counterclockwise direction.

- The point on the unit circle which corresponds to $\pi/2$ in the domain of sn is the point with coordinates $(0, 1)$.



- The point on the unit circle which corresponds to π in the domain of sn is the point with coordinates $(-1, 0)$.
- The point on the unit circle which corresponds to $3\pi/2$ in the domain of sn is the point with coordinates $(-1, -1)$.
- The point on the unit circle which corresponds to 2π in the domain of sn is the point with coordinates $(1, 0)$.

Whether in the domain of sn or on the circumference of the unit circle, these four points are at the starting point, $\frac{1}{4}$ the total distance, $\frac{1}{2}$ the total distance, and $\frac{3}{4}$ the total distance.

We are now ready to provide the rule for the functions named sn and cs.

RULE: For any $x \in [0, 2\pi]$, sn(x) is the second coordinate of the corresponding point on the circumference of the unit circle.

RULE: For any $x \in [0, 2\pi]$, cs(x) is the first coordinate of the corresponding point on the circumference of the unit circle.

The above diagram shows that:

$$\text{sn}(0) = 0, \quad \text{sn}\left(\frac{\pi}{2}\right) = 1, \quad \text{sn}(\pi) = 0, \quad \text{sn}\left(\frac{3\pi}{2}\right) = -1, \quad \text{sn}(2\pi) = 0$$

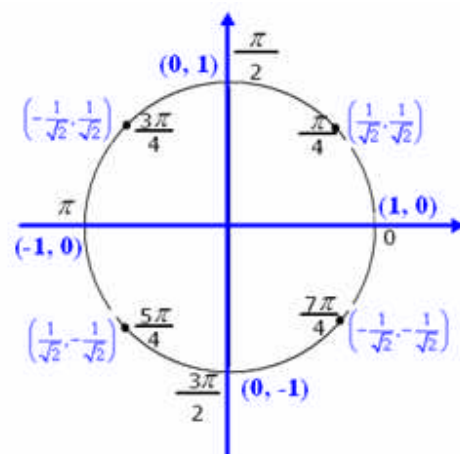
$$\text{cs}(0) = 1, \quad \text{cs}\left(\frac{\pi}{2}\right) = 0, \quad \text{cs}(\pi) = -1, \quad \text{cs}\left(\frac{3\pi}{2}\right) = 0, \quad \text{cs}(2\pi) = 1$$

We will now look at the range values associated with a few other domain elements. In particular we will examine those numbers (domain elements) midway between each pair of the previous four numbers in the domain of sn and cs.

Recall the rules for the functions sn and cs and extract the following range values directly from the picture at the right.

$$\text{sn}\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}, \quad \text{sn}\left(\frac{3\pi}{4}\right) = \frac{1}{\sqrt{2}}, \quad \text{sn}\left(\frac{5\pi}{4}\right) = -\frac{1}{\sqrt{2}}, \quad \text{sn}\left(\frac{7\pi}{4}\right) = -\frac{1}{\sqrt{2}}$$

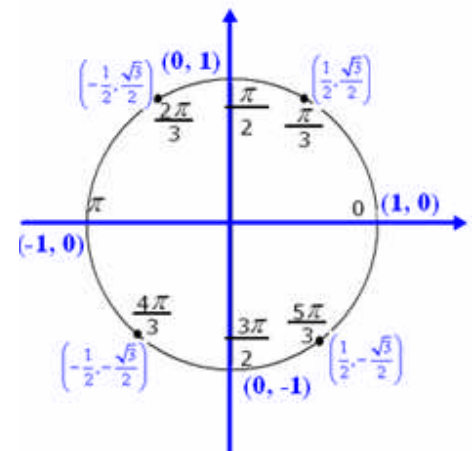
$$\text{cs}\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}, \quad \text{cs}\left(\frac{3\pi}{4}\right) = -\frac{1}{\sqrt{2}}, \quad \text{cs}\left(\frac{5\pi}{4}\right) = \frac{1}{\sqrt{2}}, \quad \text{cs}\left(\frac{7\pi}{4}\right) = -\frac{1}{\sqrt{2}}$$



The figure at the right shows additional points in the domain of sn and cs with their coordinates on the unit circle. From this diagram and the rules for the two functions we can conclude:

$$\text{sn}\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}, \quad \text{sn}\left(\frac{2\pi}{3}\right) = \frac{\sqrt{3}}{2}, \quad \text{sn}\left(\frac{4\pi}{3}\right) = -\frac{\sqrt{3}}{2}, \quad \text{sn}\left(\frac{5\pi}{3}\right) = -\frac{\sqrt{3}}{2}$$

$$\text{cs}\left(\frac{\pi}{3}\right) = \frac{1}{2}, \quad \text{cs}\left(\frac{2\pi}{3}\right) = -\frac{1}{2}, \quad \text{cs}\left(\frac{4\pi}{3}\right) = -\frac{1}{2}, \quad \text{cs}\left(\frac{5\pi}{3}\right) = \frac{1}{2}$$



In a manner similar to the examples presented in these few examples, the unique range values associated with a domain element may be determined.

Each point in the domain of sn and cs corresponds with a point on the unit circle which in turn corresponds with a set of first and second coordinates which determine the unique range value associated with the domain element.

The two functions sn and cs and four other circular functions are the functions studied in Trigonometry. How these functions relate to angles, triangles, radian measure, etc. will not be discussed here. The purpose here is simply to give an illustration of some functions whose rules are unusual and whose domains and ranges are not all of \mathbf{R} .

CS 441 Discrete Mathematics for CS
Lecture 9

Functions II

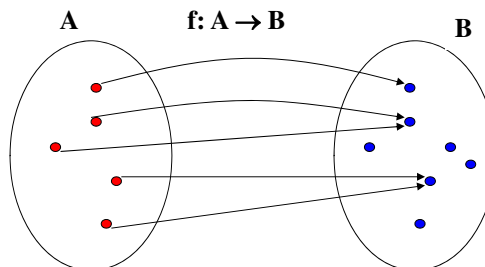
Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

CS 441 Discrete mathematics for CS

M. Hauskrecht

Functions

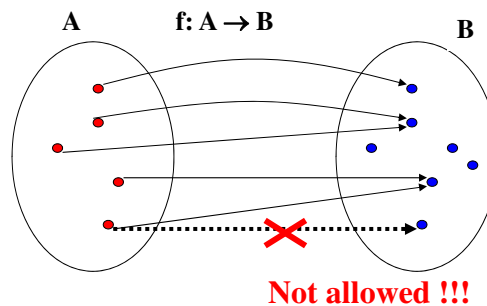
- **Definition:** Let A and B be two sets. A **function from A to B** , denoted $f: A \rightarrow B$, is an assignment of exactly one element of B to each element of A . We write $f(a) = b$ to denote the assignment of b to an element a of A by the function f .



M. Hauskrecht

Functions

- Definition:** Let A and B be two sets. A **function from A to B** , denoted $f: A \rightarrow B$, is an assignment of exactly one element of B to each element of A . We write $f(a) = b$ to denote the assignment of b to an element a of A by the function f .



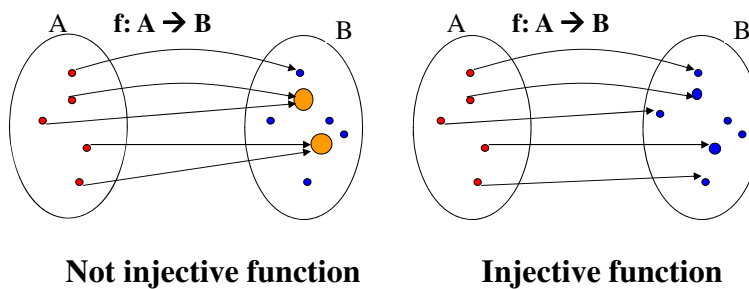
CS 441 Discrete mathematics for CS

M. Hauskrecht

Injective function

Definition: A function f is said to be **one-to-one, or injective**, if and only if $f(x) = f(y)$ implies $x = y$ for all x, y in the domain of f . A function is said to be an **injection if it is one-to-one**.

Alternative: A function is one-to-one if and only if $f(x) \neq f(y)$, whenever $x \neq y$. This is the contrapositive of the definition.



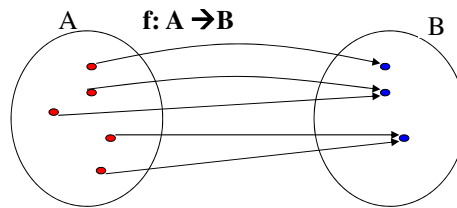
CS 441 Discrete mathematics for CS

M. Hauskrecht

Surjective function

Definition: A function f from A to B is called **onto, or surjective**, if and only if for every $b \in B$ there is an element $a \in A$ such that $f(a) = b$.

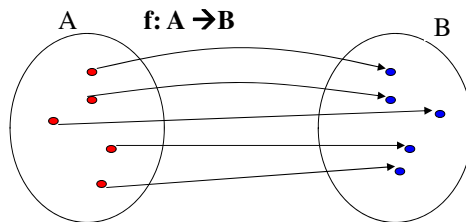
Alternative: all co-domain elements are covered



M. Hauskrecht

Bijjective functions

Definition: A function f is called **a bijection** if it is **both one-to-one (injection) and onto (surjection)**.



M. Hauskrecht

Bijjective functions

Example 1:

- Let $A = \{1,2,3\}$ and $B = \{a,b,c\}$
 - Define f as
 - $1 \rightarrow c$
 - $2 \rightarrow a$
 - $3 \rightarrow b$
- Is f a bijection?
- ?

M. Hauskrecht

Bijjective functions

Example 1:

- Let $A = \{1,2,3\}$ and $B = \{a,b,c\}$
 - Define f as
 - $1 \rightarrow c$
 - $2 \rightarrow a$
 - $3 \rightarrow b$
- Is f a bijection?
- **Yes.** It is both one-to-one and onto.

M. Hauskrecht

Bijjective functions

Example 2:

- Define $g : W \rightarrow W$ (whole numbers), where $g(n) = \lfloor n/2 \rfloor$ (floor function).
 - $0 \rightarrow \lfloor 0/2 \rfloor = \lfloor 0 \rfloor = 0$
 - $1 \rightarrow \lfloor 1/2 \rfloor = \lfloor 1/2 \rfloor = 0$
 - $2 \rightarrow \lfloor 2/2 \rfloor = \lfloor 1 \rfloor = 1$
 - $3 \rightarrow \lfloor 3/2 \rfloor = \lfloor 3/2 \rfloor = 1$
- ...
- Is g a bijection?

M. Hauskrecht

Bijjective functions

Example 2:

- Define $g : W \rightarrow W$ (whole numbers), where $g(n) = \lfloor n/2 \rfloor$ (floor function).
 - $0 \rightarrow \lfloor 0/2 \rfloor = \lfloor 0 \rfloor = 0$
 - $1 \rightarrow \lfloor 1/2 \rfloor = \lfloor 1/2 \rfloor = 0$
 - $2 \rightarrow \lfloor 2/2 \rfloor = \lfloor 1 \rfloor = 1$
 - $3 \rightarrow \lfloor 3/2 \rfloor = \lfloor 3/2 \rfloor = 1$
- ...
- Is g a bijection?
 - **No.** g is onto but not 1-1 ($g(0) = g(1) = 0$ however $0 \neq 1$).

M. Hauskrecht

Bijjective functions

Theorem: Let f be a function $f: A \rightarrow A$ from a set A to itself, where A is finite. Then f is one-to-one if and only if f is onto.

Assume

→ A is finite and f is one-to-one (injective)

- Is f an onto function (surjection)?

Bijjective functions

Theorem: Let f be a function $f: A \rightarrow A$ from a set A to itself, where A is finite. Then f is one-to-one if and only if f is onto.

Proof:

→ A is finite and f is one-to-one (injective)

- Is f an onto function (surjection)?
- **Yes.** Every element points to exactly one element. Injection assures they are different. So we have $|A|$ different elements A points to. Since $f: A \rightarrow A$ the co-domain is covered thus the function is also a surjection (and a bijection)

← A is finite and f is an onto function

- Is the function one-to-one?

Bijjective functions

Theorem: Let f be a function $f: A \rightarrow A$ from a set A to itself, where A is finite. Then f is one-to-one if and only if f is onto.

Proof:

→ **A is finite and f is one-to-one (injective)**

- Is f an onto function (surjection)?
- **Yes.** Every element points to exactly one element. Injection assures they are different. So we have $|A|$ different elements A points to. Since $f: A \rightarrow A$ the co-domain is covered thus the function is also a surjection (and a bijection)

← **A is finite and f is an onto function**

- Is the function one-to-one?
- **Yes.** Every element maps to exactly one element and all elements in A are covered. Thus the mapping must be one-to-

one

CS 441 Discrete mathematics for CS

M. Hauskrecht

Bijjective functions

Theorem. Let f be a function from a set A to itself, where A is finite. Then f is one-to-one if and only if f is onto.

Please note the above is not true when A is an infinite set.

• **Example:**

- $f: \mathbb{Z} \rightarrow \mathbb{Z}$, where $f(z) = 2 * z$.
- f is one-to-one but not onto.
 - $1 \rightarrow 2$
 - $2 \rightarrow 4$
 - $3 \rightarrow 6$
- 3 has no pre-image.

M. Hauskrecht

Functions on real numbers

Definition: Let f_1 and f_2 be functions from A to \mathbf{R} (reals). Then $f_1 + f_2$ and $f_1 * f_2$ are also functions from A to \mathbf{R} defined by

- $(f_1 + f_2)(x) = f_1(x) + f_2(x)$
- $(f_1 * f_2)(x) = f_1(x) * f_2(x).$

Examples:

- **Assume**

- $f_1(x) = x - 1$
- $f_2(x) = x^3 + 1$

then

- $(f_1 + f_2)(x) = x^3 + x$
- $(f_1 * f_2)(x) = x^4 - x^3 + x - 1.$

Increasing and decreasing functions

Definition: A function f whose domain and codomain are subsets of real numbers is **strictly increasing** if $f(x) > f(y)$ whenever $x > y$ and x and y are in the domain of f . Similarly, f is called **strictly decreasing** if $f(x) < f(y)$ whenever $x > y$ and x and y are in the domain of f .

Example:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$. Is it increasing ?

Increasing and decreasing functions

Definition: A function f whose domain and codomain are subsets of real numbers is **strictly increasing** if $f(x) > f(y)$ whenever $x > y$ and x and y are in the domain of f . Similarly, f is called **strictly decreasing** if $f(x) < f(y)$ whenever $x > y$ and x and y are in the domain of f .

Example:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$. Is it increasing ?
- **Proof .**
For $x > y$ holds $2x > 2y$ and subsequently $2x-1 > 2y-1$
Thus g is strictly increasing.

Increasing and decreasing functions

Definition: A function f whose domain and codomain are subsets of real numbers is **strictly increasing** if $f(x) > f(y)$ whenever $x > y$ and x and y are in the domain of f . Similarly, f is called **strictly decreasing** if $f(x) < f(y)$ whenever $x > y$ and x and y are in the domain of f .

Note: Strictly increasing and strictly decreasing functions are one-to-one.

Why?

Increasing and decreasing functions

Definition: A function f whose domain and codomain are subsets of real numbers is **strictly increasing** if $f(x) > f(y)$ whenever $x > y$ and x and y are in the domain of f . Similarly, f is called **strictly decreasing** if $f(x) < f(y)$ whenever $x > y$ and x and y are in the domain of f .

Note: Strictly increasing and strictly decreasing functions are one-to-one.

Why?

One-to-one function: A function is one-to-one if and only if $f(x) \neq f(y)$, whenever $x \neq y$.

Identity function

Definition: Let A be a set. The **identity function** on A is the function $i_A: A \rightarrow A$ where $i_A(x) = x$.

Example:

- Let $A = \{1, 2, 3\}$

Then:

- $i_A(1) = ?$

Identity function

Definition: Let A be a set. The **identity function** on A is the function $i_A: A \rightarrow A$ where $i_A(x) = x$.

Example:

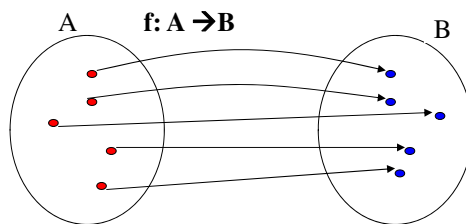
- Let $A = \{1, 2, 3\}$

Then:

- $i_A(1) = 1$
- $i_A(2) = 2$
- $i_A(3) = 3$.

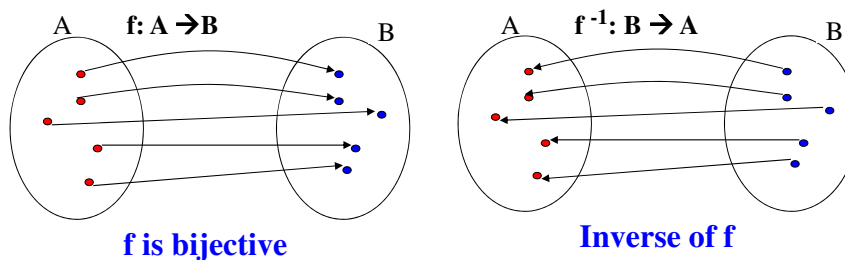
Bijjective functions

Definition: A function f is called **a bijection** if it is **both one-to-one and onto**.



Inverse functions

Definition: Let f be a **bijection** from set A to set B . The **inverse function of f** is the function that assigns to an element b from B the unique element a in A such that $f(a) = b$. The inverse function of f is denoted by f^{-1} . Hence, $f^{-1}(b) = a$, when $f(a) = b$. If the inverse function of f exists, f is called **invertible**.



CS 441 Discrete mathematics for CS

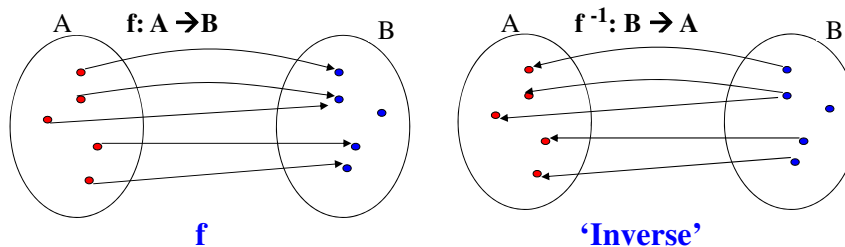
M. Hauskrecht

Inverse functions

Note: if f is not a bijection then it is not possible to define the inverse function of f . **Why?**

Assume f is not one-to-one:

?



CS 441 Discrete mathematics for CS

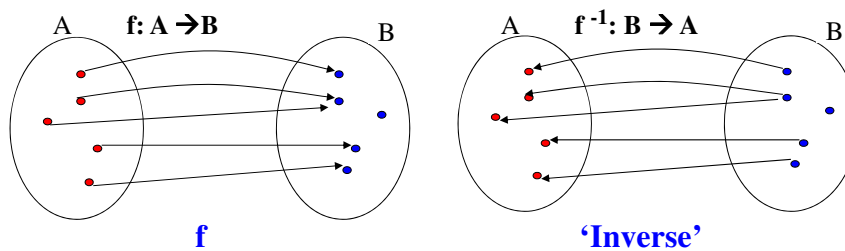
M. Hauskrecht

Inverse functions

Note: if f is not a bijection then it is not possible to define the inverse function of f . **Why?**

Assume f is not one-to-one:

Inverse is not a function. One element of B is mapped to two different elements.



CS 441 Discrete mathematics for CS

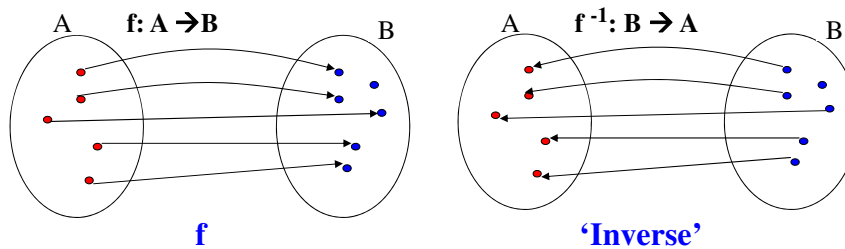
M. Hauskrecht

Inverse functions

Note: if f is not a bijection then it is not possible to define the inverse function of f . **Why?**

Assume f is not onto:

?



CS 441 Discrete mathematics for CS

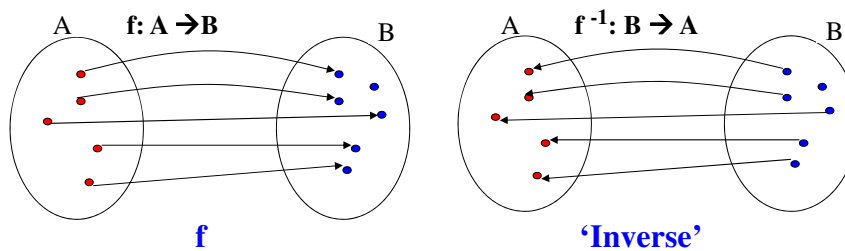
M. Hauskrecht

Inverse functions

Note: if f is not a bijection then it is not possible to define the inverse function of f . Why?

Assume f is not onto:

Inverse is not a function. One element of B is not assigned any value in B .



CS 441 Discrete mathematics for CS

M. Hauskrecht

Inverse functions

Example 1:

- Let $A = \{1, 2, 3\}$ and i_A be the identity function

- | | | |
|---|--------------|-------------------|
| • | $i_A(1) = 1$ | $i_A^{-1}(1) = 1$ |
| • | $i_A(2) = 2$ | $i_A^{-1}(2) = 2$ |
| • | $i_A(3) = 3$ | $i_A^{-1}(3) = 3$ |

- Therefore, the inverse function of i_A is i_A .

CS 441 Discrete mathematics for CS

M. Hauskrecht

Inverse functions

Example 2:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$.
- What is the inverse function g^{-1} ?

Inverse functions

Example 2:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$.
- What is the inverse function g^{-1} ?

Approach to determine the inverse:

$$y = 2x - 1 \Rightarrow y + 1 = 2x$$

$$\Rightarrow (y+1)/2 = x$$

- Define $g^{-1}(y) = x = (y+1)/2$

Test the correctness of inverse:

- $g(3) = ..$

Inverse functions

Example 2:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$.
- What is the inverse function g^{-1} ?

Approach to determine the inverse:

$$y = 2x - 1 \Rightarrow y + 1 = 2x$$

$$\Rightarrow (y+1)/2 = x$$

- Define $g^{-1}(y) = x = (y+1)/2$

Test the correctness of inverse:

- $g(3) = 2 \cdot 3 - 1 = 5$
- $g^{-1}(5) =$

Inverse functions

Example 2:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$.
- What is the inverse function g^{-1} ?

Approach to determine the inverse:

$$y = 2x - 1 \Rightarrow y + 1 = 2x$$

$$\Rightarrow (y+1)/2 = x$$

- Define $g^{-1}(y) = x = (y+1)/2$

Test the correctness of inverse:

- $g(3) = 2 \cdot 3 - 1 = 5$
- $g^{-1}(5) = (5+1)/2 = 3$
- $g(10) =$

Inverse functions

Example 2:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$.
- What is the inverse function g^{-1} ?

Approach to determine the inverse:

$$y = 2x - 1 \Rightarrow y + 1 = 2x$$

$$\Rightarrow (y+1)/2 = x$$

- Define $g^{-1}(y) = x = (y+1)/2$

Test the correctness of inverse:

- $g(3) = 2*3 - 1 = 5$
- $g^{-1}(5) = (5+1)/2 = 3$
- $g(10) = 2*10 - 1 = 19$
- $g^{-1}(19) =$

Inverse functions

Example 2:

- Let $g : \mathbf{R} \rightarrow \mathbf{R}$, where $g(x) = 2x - 1$.
- What is the inverse function g^{-1} ?

Approach to determine the inverse:

$$y = 2x - 1 \Rightarrow y + 1 = 2x$$

$$\Rightarrow (y+1)/2 = x$$

- Define $g^{-1}(y) = x = (y+1)/2$

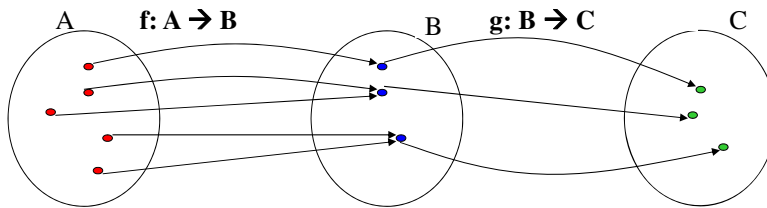
Test the correctness of inverse:

- $g(3) = 2*3 - 1 = 5$
- $g^{-1}(5) = (5+1)/2 = 3$
- $g(10) = 2*10 - 1 = 19$
- $g^{-1}(19) = (19+1)/2 = 10$.

Composition of functions

Definition: Let f be a function from set A to set B and let g be a function from set B to set C . The **composition of the functions g and f** , denoted by $g \circ f$ is defined by

- $(g \circ f)(a) = g(f(a))$.



Composition of functions

Example 1:

- Let $A = \{1,2,3\}$ and $B = \{a,b,c,d\}$

$$g : A \rightarrow A, \quad f: A \rightarrow B$$

$$1 \rightarrow 3 \quad 1 \rightarrow b$$

$$2 \rightarrow 1 \quad 2 \rightarrow a$$

$$3 \rightarrow 2 \quad 3 \rightarrow d$$

$$f \circ g : A \rightarrow B:$$

- $1 \rightarrow$

Composition of functions

Example 1:

- Let $A = \{1,2,3\}$ and $B = \{a,b,c,d\}$

$$g : A \rightarrow A, \quad f: A \rightarrow B$$

$$1 \rightarrow 3 \quad 1 \rightarrow b$$

$$2 \rightarrow 1 \quad 2 \rightarrow a$$

$$3 \rightarrow 2 \quad 3 \rightarrow d$$

$$f \circ g : A \rightarrow B:$$

- $1 \rightarrow d$
- $2 \rightarrow$

Composition of functions

Example 1:

- Let $A = \{1,2,3\}$ and $B = \{a,b,c,d\}$

$$g : A \rightarrow A, \quad f: A \rightarrow B$$

$$1 \rightarrow 3 \quad 1 \rightarrow b$$

$$2 \rightarrow 1 \quad 2 \rightarrow a$$

$$3 \rightarrow 2 \quad 3 \rightarrow d$$

$$f \circ g : A \rightarrow B:$$

- $1 \rightarrow d$
- $2 \rightarrow b$
- $3 \rightarrow$

Composition of functions

Example 1:

- Let $A = \{1,2,3\}$ and $B = \{a,b,c,d\}$

$$g : A \rightarrow A, \quad f: A \rightarrow B$$

$$1 \rightarrow 3 \quad 1 \rightarrow b$$

$$2 \rightarrow 1 \quad 2 \rightarrow a$$

$$3 \rightarrow 2 \quad 3 \rightarrow d$$

$$f \circ g : A \rightarrow B:$$

- $1 \rightarrow d$
- $2 \rightarrow b$
- $3 \rightarrow a$

Composition of functions

Example 2:

- Let f and g be two functions from Z to Z , where
- $f(x) = 2x$ and $g(x) = x^2$.
- $f \circ g : Z \rightarrow Z$
- $$\begin{aligned} (f \circ g)(x) &= f(g(x)) \\ &= f(x^2) \\ &= 2(x^2) \end{aligned}$$
- $g \circ f : Z \rightarrow Z$
- $(g \circ f)(x) = ?$

Composition of functions

Example 2:

- Let f and g be two functions from Z to Z , where
 - $f(x) = 2x$ and $g(x) = x^2$.
 - $f \circ g : Z \rightarrow Z$
 - $$\begin{aligned}(f \circ g)(x) &= f(g(x)) \\ &= f(x^2) \\ &= 2(x^2)\end{aligned}$$
 - $g \circ f : Z \rightarrow Z$
 - $$\begin{aligned}(g \circ f)(x) &= g(f(x)) \\ &= g(2x) \\ &= (2x)^2 \\ &= 4x^2\end{aligned}$$
- Note that the order of the function composition matters

Composition of functions

Example 3:

- $(f \circ f^{-1})(x) = x$ and $(f^{-1} \circ f)(x) = x$, for all x .
- Let $f : \mathbf{R} \rightarrow \mathbf{R}$, where $f(x) = 2x - 1$ and $f^{-1}(x) = (x+1)/2$.
- $$\begin{aligned}(f \circ f^{-1})(x) &= f(f^{-1}(x)) \\ &= f((x+1)/2) \\ &= 2((x+1)/2) - 1 \\ &= (x+1) - 1 \\ &= x\end{aligned}$$

Composition of functions

Example 3:

- $(f \circ f^{-1})(x) = x$ and $(f^{-1} \circ f)(x) = x$, for all x .
- Let $f : \mathbf{R} \rightarrow \mathbf{R}$, where $f(x) = 2x - 1$ and $f^{-1}(x) = (x+1)/2$.
- $(f \circ f^{-1})(x) = f(f^{-1}(x))$
 $= f((x+1)/2)$
 $= 2((x+1)/2) - 1$
 $= (x+1) - 1$
 $= x$
- $(f^{-1} \circ f)(x) = f^{-1}(f(x))$
 $= f^{-1}(2x - 1)$
 $= (2x)/2$
 $= x$

Some functions

Definitions:

- The **floor function** assigns a real number x the largest integer that is less than or equal to x . The floor function is denoted by $\lfloor x \rfloor$.
- The **ceiling function** assigns to the real number x the smallest integer that is greater than or equal to x . The ceiling function is denoted by $\lceil x \rceil$.

Other important functions:

- Factorials: $n! = n(n-1) \dots 1$ such that $1! = 1$

Chapter three:

Functions and their graphs

1 Functions

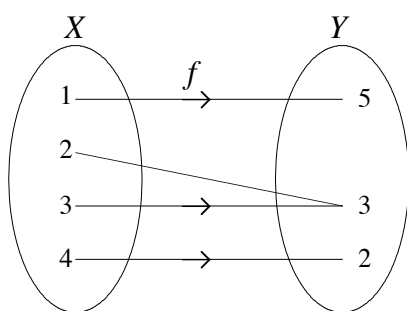
In this Chapter we will cover various aspects of functions. We will look at the definition of a function, the domain and range of a function, what we mean by specifying the domain of a function and absolute value function.

1.1 What is a function?

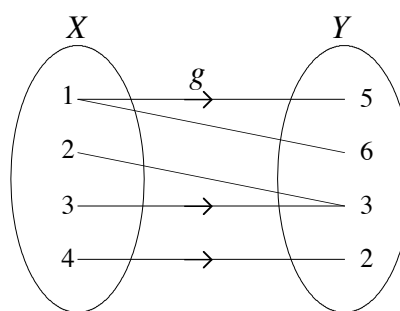
1.1.1 Definition of a function

A function f from a set of elements X to a set of elements Y is a rule that assigns to each element x in X exactly one element y in Y .

One way to demonstrate the meaning of this definition is by using arrow diagrams.



$f : X \rightarrow Y$ is a function. Every element in X has associated with it exactly one element of Y .



$g : X \rightarrow Y$ is not a function. The element 1 in set X is assigned two elements, 5 and 6 in set Y .

A function can also be described as a set of ordered pairs (x, y) such that for any x -value in the set, there is only one y -value. This means that there cannot be any repeated x -values with different y -values.

The examples above can be described by the following sets of ordered pairs.

$F = \{(1,5), (3,3), (2,3), (4,2)\}$ is a function.

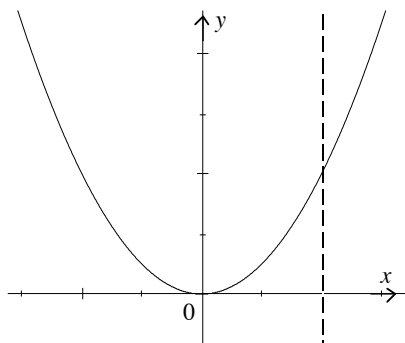
$G = \{(1,5), (4,2), (2,3), (3,3), (1,6)\}$ is not a function.

The definition we have given is a general one. While in the examples we have used numbers as elements of X and Y , there is no reason why this must be so. However, in these notes we will only consider functions where X and Y are subsets of the real numbers.

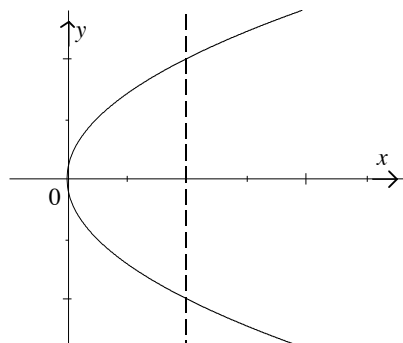
In this setting, we often describe a function using the rule, $y = f(x)$, and create a graph of that function by plotting the ordered pairs $(x, f(x))$ on the Cartesian Plane. This graphical representation allows us to use a test to decide whether or not we have the graph of a function: The Vertical Line Test.

1.1.2 The Vertical Line Test

The Vertical Line Test states that if it is *not possible* to draw a vertical line through a graph so that it cuts the graph in more than one point, then the graph *is* a function.



This is the graph of a function. All possible vertical lines will cut this graph only once.



This is not the graph of a function. The vertical line we have drawn cuts the graph twice.

1.1.3 Domain of a function

For a function $f : X \rightarrow Y$ the *domain* of f is the set X .

This also corresponds to the set of x -values when we describe a function as a set of ordered pairs (x, y) .

If only the rule $y = f(x)$ is given, then the domain is taken to be the set of all real x for which the function is defined. For example, $y = \sqrt{x}$ has domain; all real $x \geq 0$. This is sometimes referred to as the *natural* domain of the function.

1.1.4 Range of a function

For a function $f : X \rightarrow Y$ the *range* of f is the set of y -values such that $y = f(x)$ for some x in X .

This corresponds to the set of y -values when we describe a function as a set of ordered pairs (x, y) . The function $y = \sqrt{x}$ has range; all real $y \geq 0$.

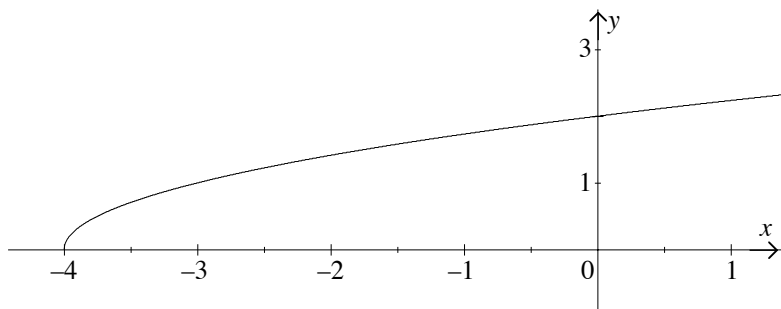
Example

- State the domain and range of $y = \sqrt{x+4}$.
- Sketch, showing significant features, the graph of $y = \sqrt{x+4}$.

Solution

a. The domain of $y = \sqrt{x+4}$ is all real $x \geq -4$. We know that square root functions are only defined for positive numbers so we require that $x+4 \geq 0$, ie $x \geq -4$. We also know that the square root functions are always positive so the range of $y = \sqrt{x+4}$ is all real $y \geq 0$.

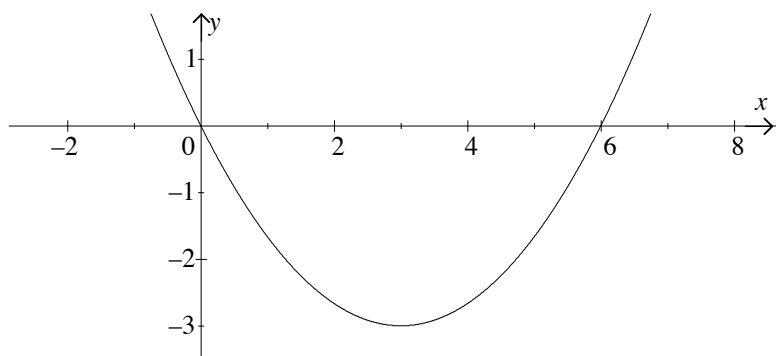
b.



The graph of $y = \sqrt{x+4}$.

Example

a. State the equation of the parabola sketched below, which has vertex $(3, -3)$.



b. Find the domain and range of this function.

Solution

a. The equation of the parabola is $y = \frac{x^2-6x}{3}$.

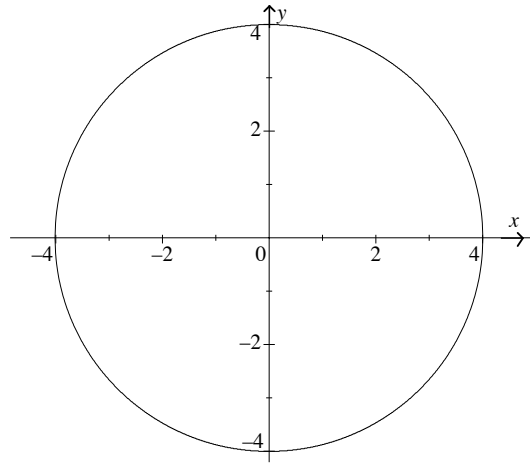
b. The domain of this parabola is all real x . The range is all real $y \geq -3$.

Example

Sketch $x^2 + y^2 = 16$ and explain why it is not the graph of a function.

Solution

$x^2 + y^2 = 16$ is not a function as it fails the vertical line test. For example, when $x = 0$ $y = \pm 4$.



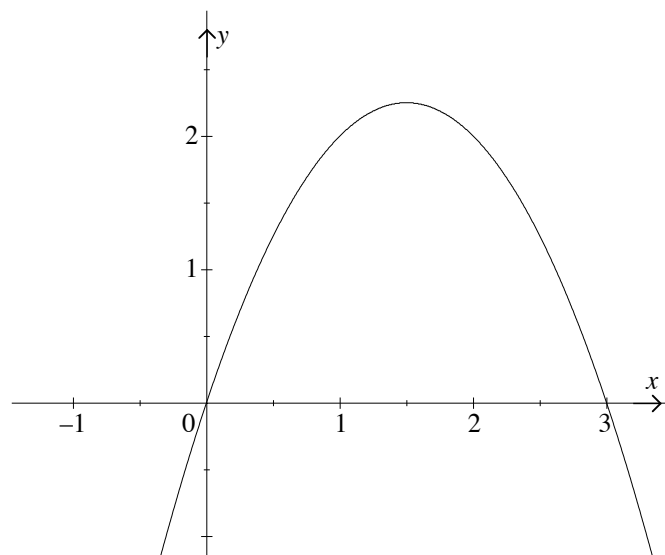
The graph of $x^2 + y^2 = 16$.

Example

Sketch the graph of $f(x) = 3x - x^2$ and find

- the domain and range
- $f(q)$
- $f(x^2)$
- $\frac{f(2+h)-f(2)}{h}$, $h \neq 0$.

Solution



The graph of $f(x) = 3x - x^2$.

- The domain is all real x . The range is all real y where $y \leq 2.25$.
- $f(q) = 3q - q^2$

c. $f(x^2) = 3(x^2) - (x^2)^2 = 3x^2 - x^4$

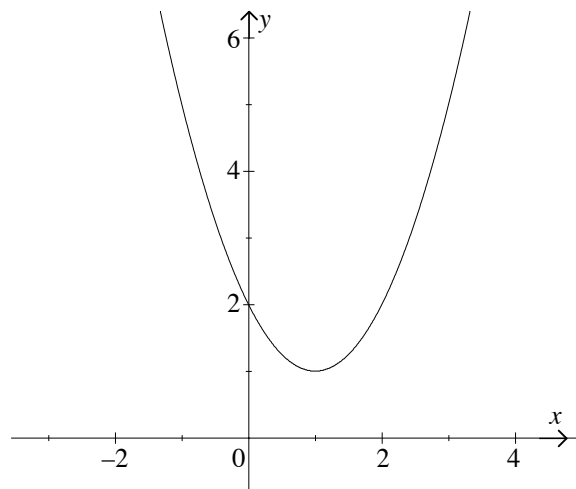
d.

$$\begin{aligned}\frac{f(2+h) - f(2)}{h} &= \frac{(3(2+h) - (2+h)^2) - (3(2) - (2)^2)}{h} \\ &= \frac{6 + 3h - (h^2 + 4h + 4) - 2}{h} \\ &= \frac{-h^2 - h}{h} \\ &= -h - 1\end{aligned}$$

Example

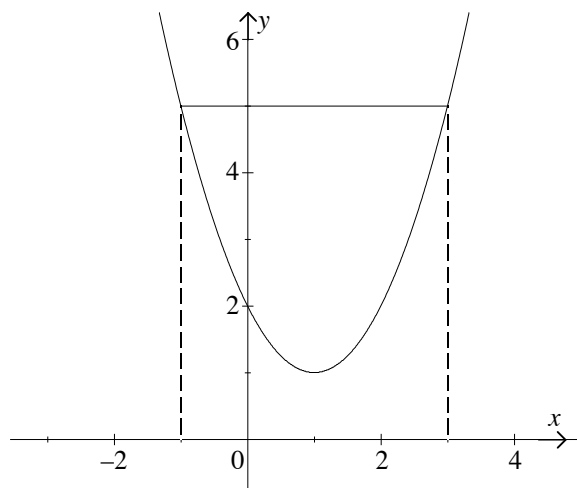
Sketch the graph of the function $f(x) = (x - 1)^2 + 1$ and show that $f(p) = f(2 - p)$. Illustrate this result on your graph by choosing one value of p .

Solution



The graph of $f(x) = (x - 1)^2 + 1$.

$$\begin{aligned}f(2 - p) &= ((2 - p) - 1)^2 + 1 \\ &= (1 - p)^2 + 1 \\ &= (p - 1)^2 + 1 \\ &= f(p)\end{aligned}$$



The sketch illustrates the relationship $f(p) = f(2 - p)$ for $p = -1$. If $p = -1$ then $2 - p = 2 - (-1) = 3$, and $f(-1) = f(3)$.

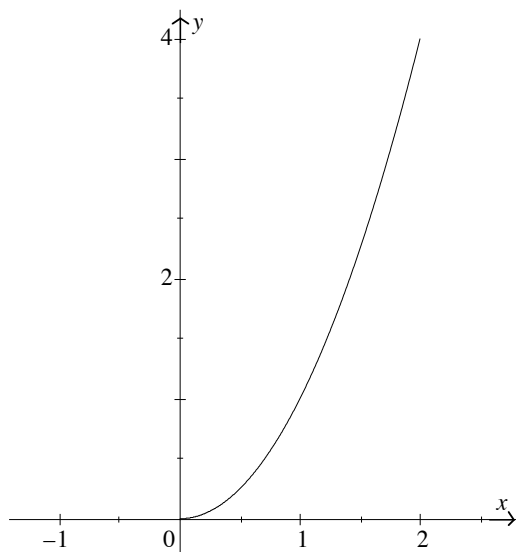
1.2 Specifying or restricting the domain of a function

We sometimes give the rule $y = f(x)$ along with the domain of definition. This domain may not necessarily be the natural domain. For example, if we have the function

$$y = x^2 \quad \text{for} \quad 0 \leq x \leq 2$$

then the domain is given as $0 \leq x \leq 2$. The natural domain has been restricted to the subinterval $0 \leq x \leq 2$.

Consequently, the range of this function is all real y where $0 \leq y \leq 4$. We can best illustrate this by sketching the graph.



The graph of $y = x^2$ for $0 \leq x \leq 2$.

1.3 The absolute value function

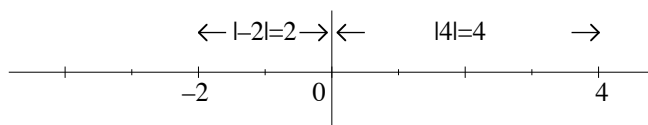
Before we define the absolute value function we will review the definition of the absolute value of a number.

The *Absolute value of a number* x is written $|x|$ and is defined as

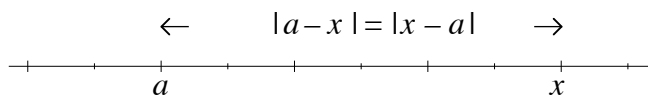
$$|x| = x \text{ if } x \geq 0 \quad \text{or} \quad |x| = -x \text{ if } x < 0.$$

That is, $|4| = 4$ since 4 is positive, but $|-2| = 2$ since -2 is negative.

We can also think of $|x|$ geometrically as the distance of x from 0 on the number line.



More generally, $|x - a|$ can be thought of as the distance of x from a on the numberline.



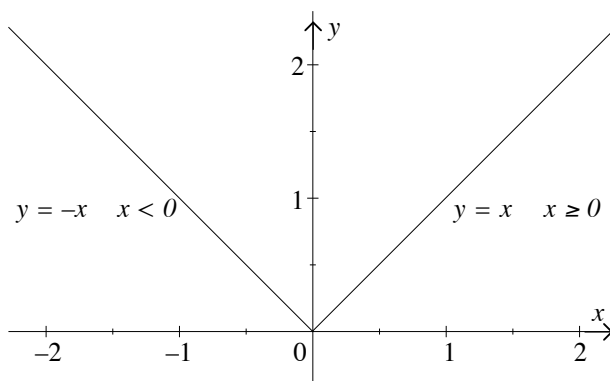
Note that $|a - x| = |x - a|$.

The absolute value *function* is written as $y = |x|$.

We define this function as

$$y = \begin{cases} +x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

From this definition we can graph the function by taking each part separately. The graph of $y = |x|$ is given below.



The graph of $y = |x|$.

Example

Sketch the graph of $y = |x - 2|$.

Solution

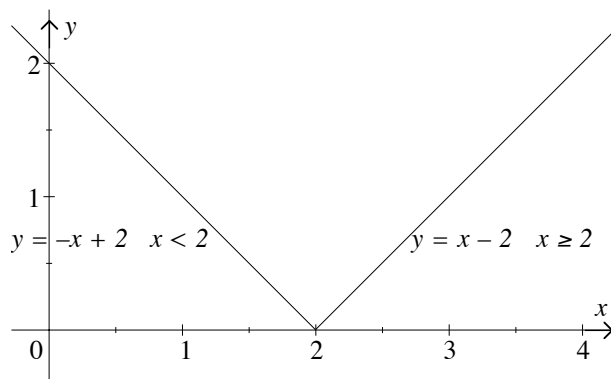
For $y = |x - 2|$ we have

$$y = \begin{cases} +(x - 2) & \text{when } x - 2 \geq 0 \quad \text{or} \quad x \geq 2 \\ -(x - 2) & \text{when } x - 2 < 0 \quad \text{or} \quad x < 2 \end{cases}$$

That is,

$$y = \begin{cases} x - 2 & \text{for } x \geq 2 \\ -x + 2 & \text{for } x < 2 \end{cases}$$

Hence we can draw the graph in two parts.



The graph of $y = |x - 2|$.

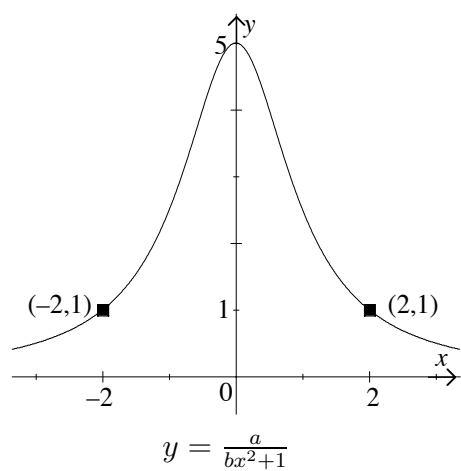
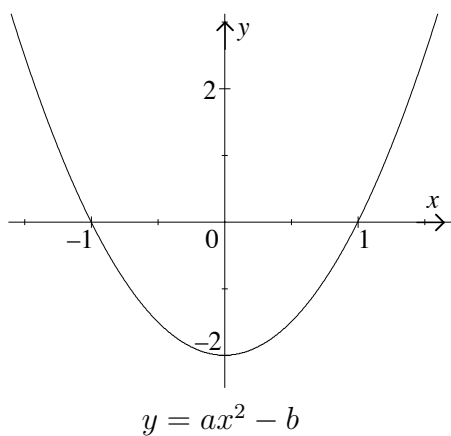
We could have sketched this graph by first of all sketching the graph of $y = x - 2$ and then reflecting the negative part in the x -axis. We will use this fact to sketch graphs of this type in Chapter 2.

1.4 Exercises

- State the domain and range of $f(x) = \sqrt{9 - x^2}$.
 - Sketch the graph of $y = \sqrt{9 - x^2}$.
- Given $\psi(x) = x^2 + 5$, find, in simplest form, $\frac{\psi(x + h) - \psi(x)}{h}$ $h \neq 0$.
- Sketch the following functions stating the domain and range of each:
 - $y = \sqrt{x - 1}$

- b. $y = |2x|$
 - c. $y = \frac{1}{x-4}$
 - d. $y = |2x| - 1$.
4. a. Find the perpendicular distance from $(0, 0)$ to the line $x + y + k = 0$
- b. If the line $x + y + k = 0$ cuts the circle $x^2 + y^2 = 4$ in two distinct points, find the restrictions on k .
5. Sketch the following, showing their important features.
- a. $y = \left(\frac{1}{2}\right)^x$
 - b. $y^2 = x^2$.
6. Explain the meanings of function, domain and range. Discuss whether or not $y^2 = x^3$ is a function.
7. Sketch the following relations, showing all intercepts and features. State which ones are functions giving their domain and range.
- a. $y = -\sqrt{4 - x^2}$
 - b. $|x| - |y| = 0$
 - c. $y = x^3$
 - d. $y = \frac{x}{|x|}, x \neq 0$
 - e. $|y| = x$.
8. If $A(x) = x^2 + 2 + \frac{1}{x^2}$, $x \neq 0$, prove that $A(p) = A\left(\frac{1}{p}\right)$ for all $p \neq 0$.
9. Write down the values of x which are not in the domain of the following functions:
- a. $f(x) = \sqrt{x^2 - 4x}$
 - b. $g(x) = \frac{x}{x^2 - 1}$
10. If $\phi(x) = \log\left(\frac{x}{x-1}\right)$, find in simplest form:
- a. $\phi(3) + \phi(4) + \phi(5)$
 - b. $\phi(3) + \phi(4) + \phi(5) + \cdots + \phi(n)$
11. a. If $y = x^2 + 2x$ and $x = (z - 2)^2$, find y when $z = 3$.
- b. Given $L(x) = 2x + 1$ and $M(x) = x^2 - x$, find
- i $L(M(x))$
 - ii $M(L(x))$

12. Using the sketches, find the value(s) of the constants in the given equations:



13. a. Define $|a|$, the absolute value of a , where a is real.

b. Sketch the relation $|x| + |y| = 1$.

14. Given that $S(n) = \frac{n}{2n+1}$, find an expression for $S(n-1)$.

Hence show that $S(n) - S(n-1) = \frac{1}{(2n-1)(2n+1)}$.

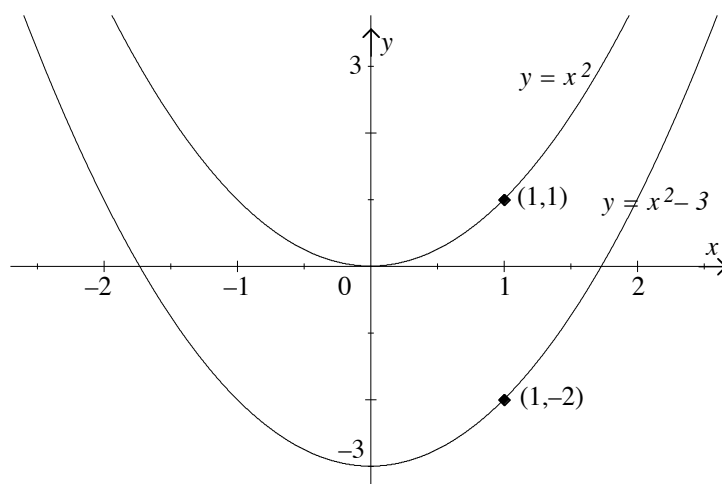
2 More about functions

In this Chapter we will look at the effects of stretching, shifting and reflecting the basic functions, $y = x^2$, $y = x^3$, $y = \frac{1}{x}$, $y = |x|$, $y = a^x$, $x^2 + y^2 = r^2$. We will introduce the concepts of even and odd functions, increasing and decreasing functions and will solve equations using graphs.

2.1 Modifying functions by shifting

2.1.1 Vertical shift

We can draw the graph of $y = f(x) + k$ from the graph of $y = f(x)$ as the addition of the constant k produces a *vertical shift*. That is, adding a constant to a function moves the graph up k units if $k > 0$ or down k units if $k < 0$. For example, we can sketch the function $y = x^2 - 3$ from our knowledge of $y = x^2$ by shifting the graph of $y = x^2$ *down* by 3 units. That is, if $f(x) = x^2$ then $f(x) - 3 = x^2 - 3$.

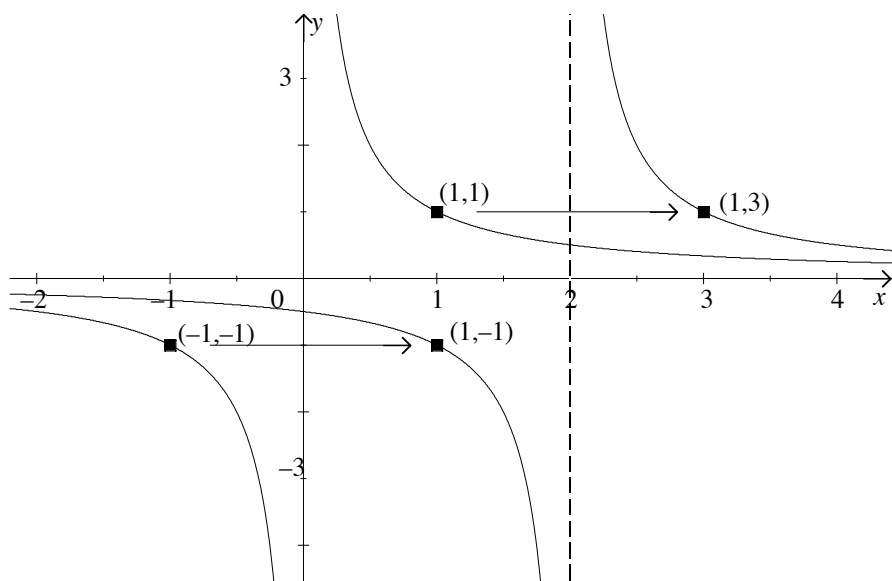


We can also write $y = f(x) - 3$ as $y + 3 = f(x)$, so replacing y by $y + 3$ in $y = f(x)$ also shifts the graph down by 3 units.

2.1.2 Horizontal shift

We can draw the graph of $y = f(x - a)$ if we know the graph of $y = f(x)$ as placing the constant a inside the brackets produces a *horizontal shift*. If we replace x by $x - a$ inside the function then the graph will shift to the left by a units if $a < 0$ and to the right by a units if $a > 0$.

For example we can sketch the graph of $y = \frac{1}{x-2}$ from our knowledge of $y = \frac{1}{x}$ by shifting this graph to the right by 2 units. That is, if $f(x) = \frac{1}{x}$ then $f(x-2) = \frac{1}{x-2}$.

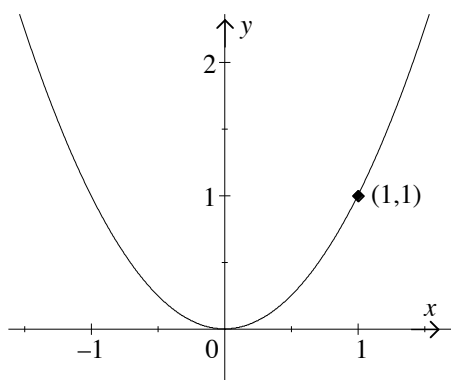


Note that the function $y = \frac{1}{x-2}$ is not defined at $x = 2$. The point $(1, 1)$ has been shifted to $(1, 3)$.

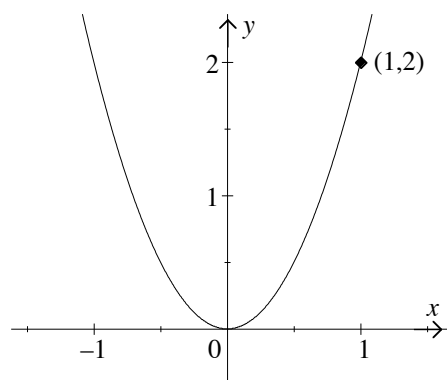
2.2 Modifying functions by stretching

We can sketch the graph of a function $y = bf(x)$ ($b > 0$) if we know the graph of $y = f(x)$ as multiplying by the constant b will have the effect of stretching the graph in the y -direction by a factor of b . That is, multiplying $f(x)$ by b will change all of the y -values proportionally.

For example, we can sketch $y = 2x^2$ from our knowledge of $y = x^2$ as follows:

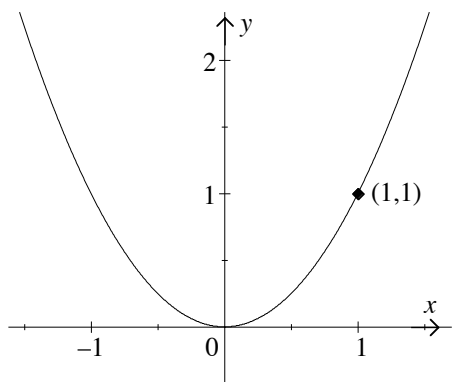


The graph of $y = x^2$.

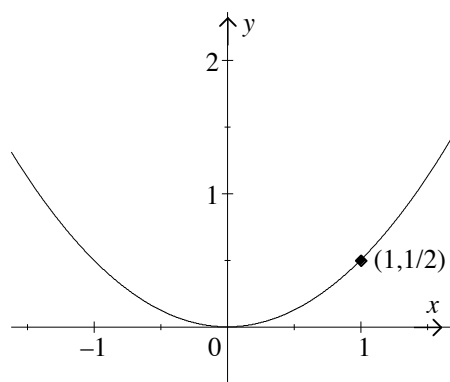


The graph of $y = 2x^2$. Note, all the y -values have been multiplied by 2, but the x -values are unchanged.

We can sketch the graph of $y = \frac{1}{2}x^2$ from our knowledge of $y = x^2$ as follows:



The graph of $y = x^2$.

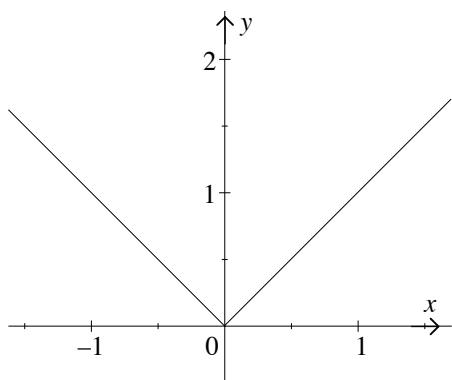


The graph of $y = \frac{1}{2}x^2$. Note, all the y -values have been multiplied by $\frac{1}{2}$, but the x -values are unchanged.

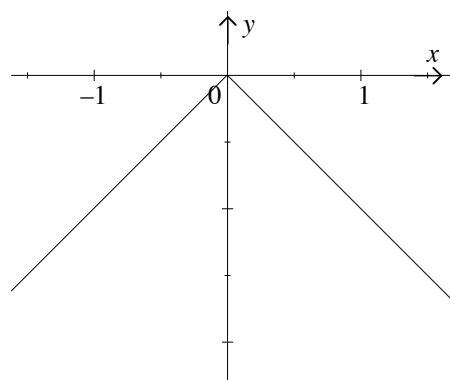
2.3 Modifying functions by reflections

2.3.1 Reflection in the x -axis

We can sketch the function $y = -f(x)$ if we know the graph of $y = f(x)$, as a minus sign in front of $f(x)$ has the effect of reflecting the whole graph in the x -axis. (Think of the x -axis as a mirror.) For example, we can sketch $y = -|x|$ from our knowledge of $y = |x|$.



The graph of $y = |x|$.

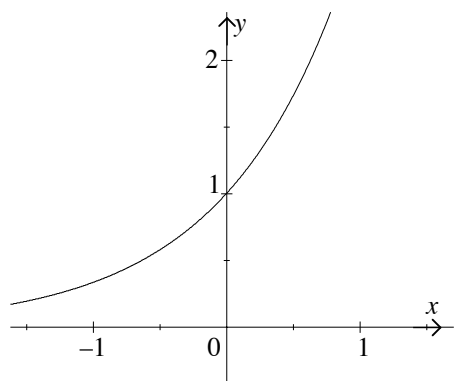


The graph of $y = -|x|$. It is the reflection of $y = |x|$ in the x -axis.

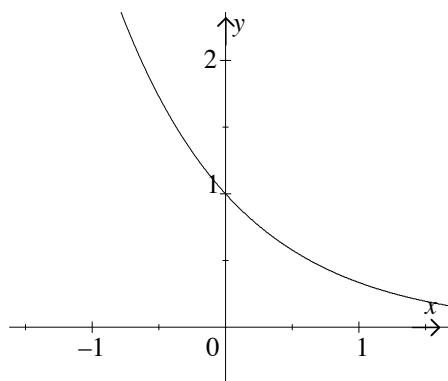
2.3.2 Reflection in the y -axis

We can sketch the graph of $y = f(-x)$ if we know the graph of $y = f(x)$ as the graph of $y = f(-x)$ is the reflection of $y = f(x)$ in the y -axis.

For example, we can sketch $y = 3^{-x}$ from our knowledge of $y = 3^x$.



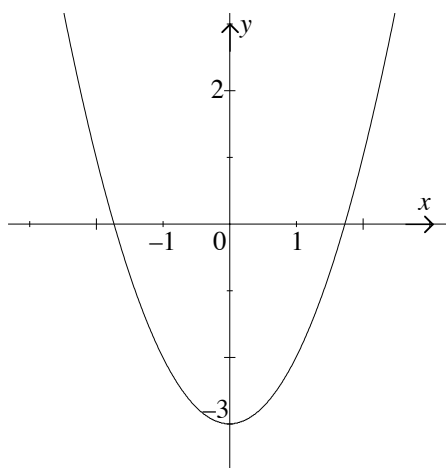
The graph of $y = 3^x$.



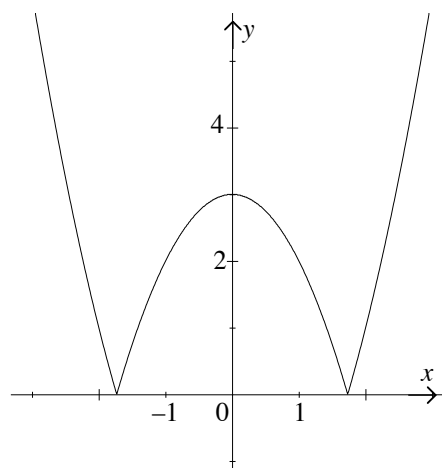
The graph of $y = 3^{-x}$. It is the reflection of $y = 3^x$ in the y -axis.

2.4 Other effects

We can sketch the graph of $y = |f(x)|$ if we know the graph of $y = f(x)$ as the effect of the absolute value is to reflect all of the *negative* values of $f(x)$ in the x -axis. For example, we can sketch the graph of $y = |x^2 - 3|$ from our knowledge of the graph of $y = x^2 - 3$.



The graph of $y = x^2 - 3$.

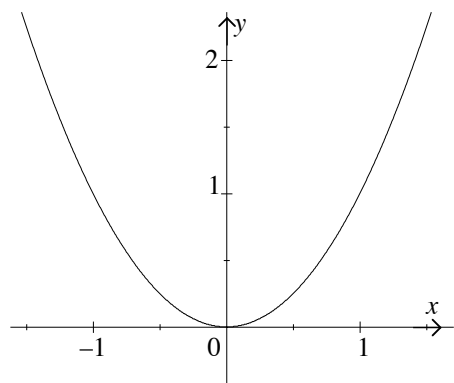


The graph of $y = |x^2 - 3|$. The negative values of $y = x^2 - 3$ have been reflected in the x -axis.

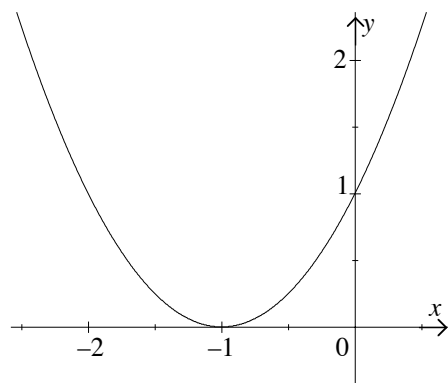
2.5 Combining effects

We can use all the above techniques to graph more complex functions. For example, we can sketch the graph of $y = 2 - (x + 1)^2$ from the graph of $y = x^2$ provided we can analyse the combined effects of the modifications. Replacing x by $x + 1$ (or $x - (-1)$) moves the

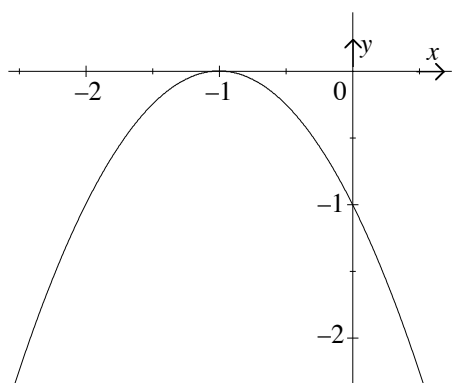
graph to the left by 1 unit. The effect of the $-$ sign in front of the brackets turns the graph up side down. The effect of adding 2 moves the graph up 2 units. We can illustrate these effects in the following diagrams.



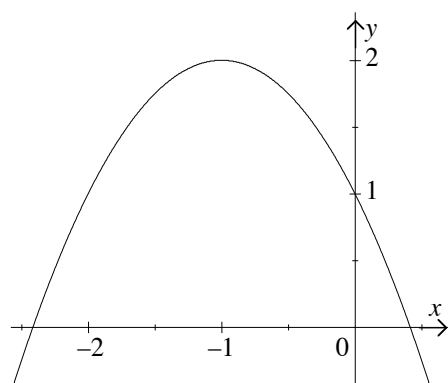
The graph of $y = x^2$.



The graph of $y = (x + 1)^2$. The graph of $y = x^2$ has been shifted 1 unit to the left.



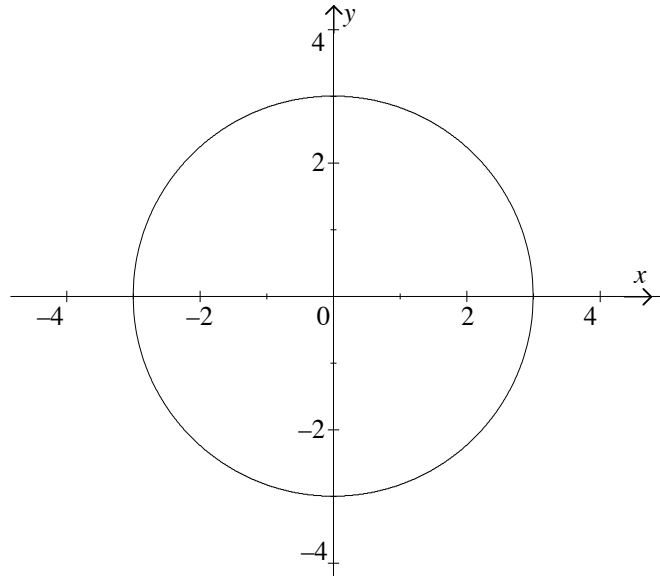
The graph of $y = -(x + 1)^2$. The graph of $y = (x + 1)^2$ has been reflected in the x -axis.



The graph of $y = 2 - (x + 1)^2$. The graph of $y = -(x + 1)^2$ has been shifted up by 2 units.

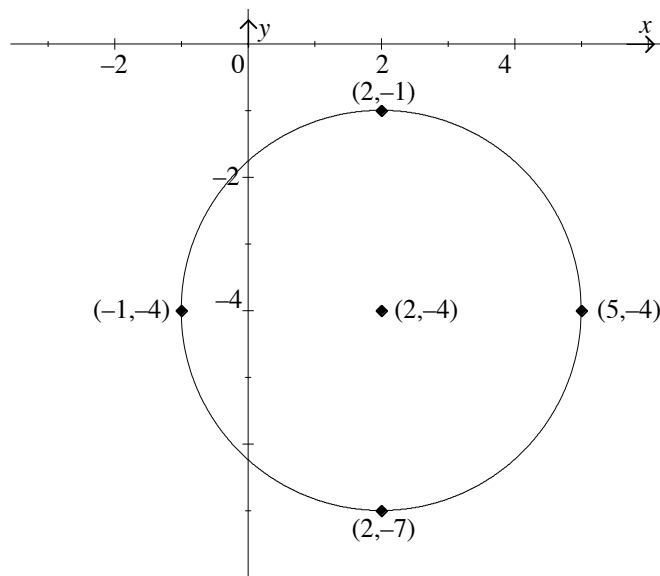
Similarly, we can sketch the graph of $(x - h)^2 + (y - k)^2 = r^2$ from the graph of $x^2 + y^2 = r^2$. Replacing x by $x - h$ shifts the graph sideways h units. Replacing y by $y - k$ shifts the graph up or down k units. (We remarked before that $y = f(x) + k$ could be written as $y - k = f(x)$.)

For example, we can use the graph of the circle of radius 3, $x^2 + y^2 = 9$, to sketch the graph of $(x - 2)^2 + (y + 4)^2 = 9$.



The graph of $x^2 + y^2 = 9$.

This is a circle centre $(0,0)$, radius 3.



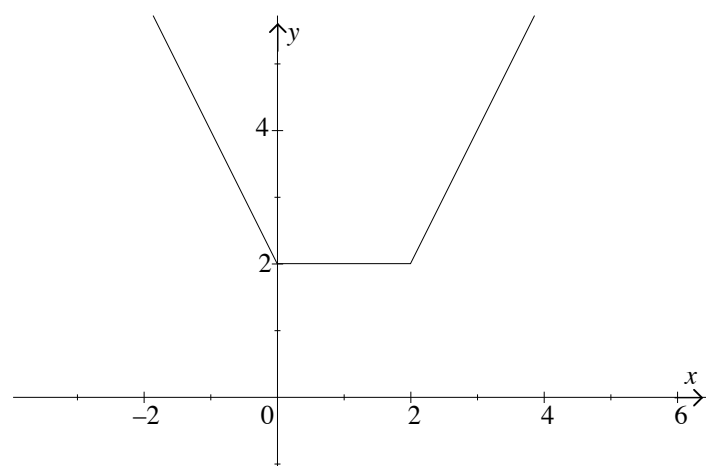
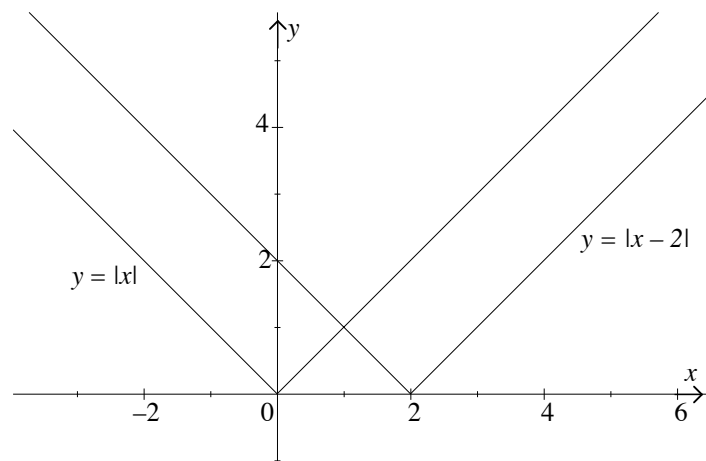
The graph of $(x - 2)^2 + (y + 4)^2 = 9$.

This is a circle centre $(2, -4)$, radius 3.

Replacing x by $x - 2$ has the effect of shifting the graph of $x^2 + y^2 = 9$ two units to the right. Replacing y by $y + 4$ shifts it down 4 units.

2.6 Graphing by addition of ordinates

We can sketch the graph of functions such as $y = |x| + |x - 2|$ by drawing the graphs of both $y = |x|$ and $y = |x - 2|$ on the same axes then adding the corresponding y -values.



The graph of $y = |x| + |x - 2|$.

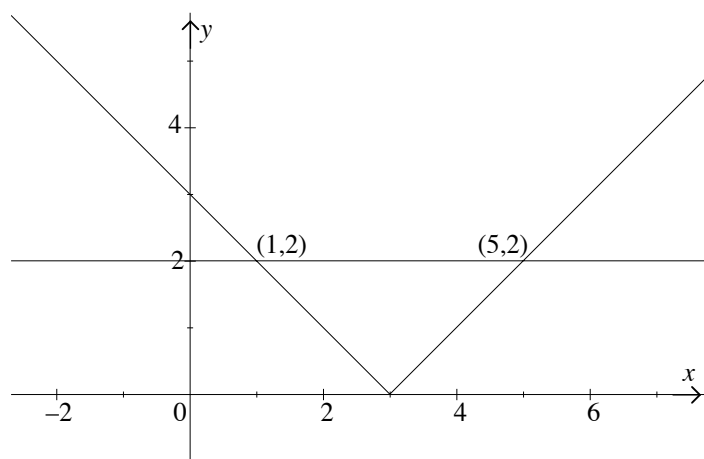
At each point of x the y -values of $y = |x|$ and $y = |x - 2|$ have been added. This allows us to sketch the graph of $y = |x| + |x - 2|$.

This technique for sketching graphs is very useful for sketching the graph of the sum of two trigonometric functions.

2.7 Using graphs to solve equations

We can solve equations of the form $f(x) = k$ by sketching $y = f(x)$ and the horizontal line $y = k$ on the same axes. The solution to the equation $f(x) = k$ is found by determining the x -values of any points of intersection of the two graphs.

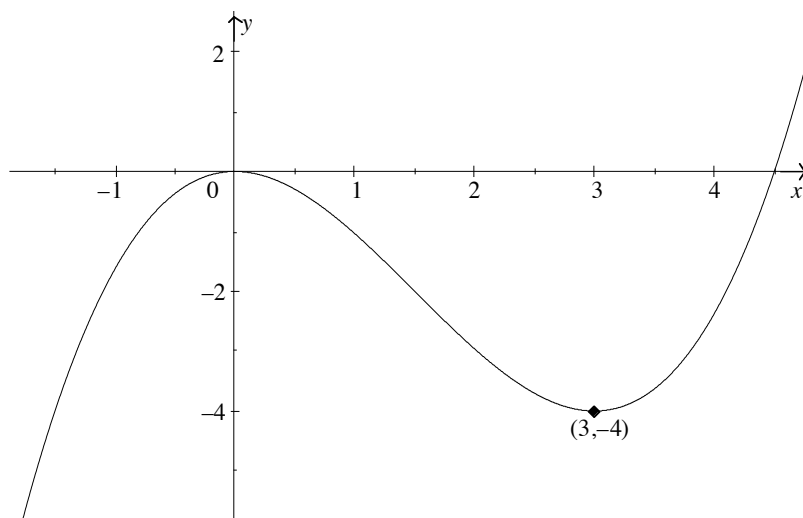
For example, to solve $|x - 3| = 2$ we sketch $y = |x - 3|$ and $y = 2$ on the same axes.



The x -values of the points of intersection are 1 and 5. Therefore $|x - 3| = 2$ when $x = 1$ or $x = 5$.

Example

The graph of $y = f(x)$ is sketched below.



For what values of k does the equation $f(x) = k$ have

1. 1 solution
2. 2 solutions
3. 3 solutions?

Solution

If we draw a horizontal line $y = k$ across the graph $y = f(x)$, it will intersect once when $k > 0$ or $k < -4$, twice when $k = 0$ or $k = -4$ and three times when $-4 < k < 0$. Therefore the equation $f(x) = k$ will have

1. 1 solution if $k > 0$ or $k < -4$
2. 2 solutions if $k = 0$ or $k = -4$
3. 3 solutions if $-4 < k < 0$.

2.8 Exercises

1. Sketch the following:

a. $y = x^2$ b. $y = \frac{1}{3}x^2$ c. $y = -x^2$ d. $y = (x + 1)^2$

2. Sketch the following:

a. $y = \frac{1}{x}$ b. $y = \frac{1}{x-2}$ c. $y = \frac{-2}{x}$ d. $y = \frac{1}{x+1} + 2$

3. Sketch the following:

a. $y = x^3$ b. $y = |x^3 - 2|$ c. $y = 3 - (x - 1)^3$

4. Sketch the following:

a. $y = |x|$ b. $y = 2|x - 2|$ c. $y = 4 - |x|$

5. Sketch the following:

a. $x^2 + y^2 = 16$ b. $x^2 + (y + 2)^2 = 16$ c. $(x - 1)^2 + (y - 3)^2 = 16$

6. Sketch the following:

a. $y = \sqrt{9 - x^2}$ b. $y = \sqrt{9 - (x - 1)^2}$ c. $y = \sqrt{9 - x^2} - 3$

7. Show that $\frac{x - 1}{x - 2} = \frac{1}{x - 2} + 1$.

Hence sketch the graph of $y = \frac{x - 1}{x - 2}$.

8. Sketch $y = \frac{x+1}{x-1}$.

9. Graph the following relations in the given interval:

a. $y = |x| + x + 1$ for $-2 \leq x \leq 2$ [Hint: Sketch by adding ordinates]

b. $y = |x| + |x - 1|$ for $-2 \leq x \leq 3$

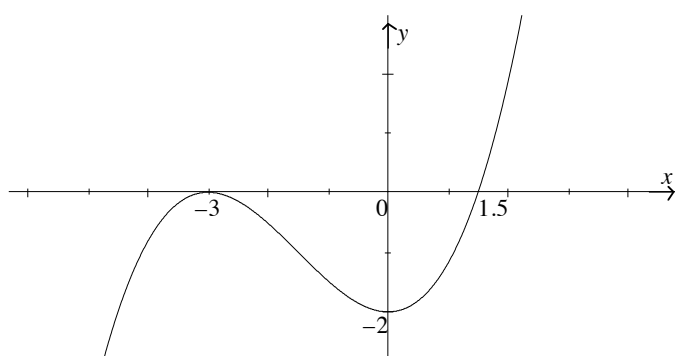
c. $y = 2^x + 2^{-x}$ for $-2 \leq x \leq 2$

d. $|x - y| = 1$ for $-1 \leq x \leq 3$.

10. Sketch the function $f(x) = |x^2 - 1| - 1$.

11. Given $y = f(x)$ as sketched below, sketch

- a. $y = 2f(x)$
- b. $y = -f(x)$
- c. $y = f(-x)$
- d. $y = f(x) + 4$
- e. $y = f(x - 3)$
- f. $y = f(x + 1) - 2$
- g. $y = 3 - 2f(x - 3)$
- h. $y = |f(x)|$



12. By sketching graphs solve the following equations:

- a. $|2x| = 4$
- b. $\frac{1}{x-2} = -1$
- c. $x^3 = x^2$
- d. $x^2 = \frac{1}{x}$

13. Solve $|x - 2| = 3$.

- a. algebraically
- b. geometrically.

14. The parabolas $y = (x - 1)^2$ and $y = (x - 3)^2$ intersect at a point P . Find the coordinates of P .

15. Sketch the circle $x^2 + y^2 - 2x - 14y + 25 = 0$. [Hint: Complete the squares.] Find the values of k , so that the line $y = k$ intersects the circle in two distinct points.

16. Solve $\frac{4}{5-x} = 1$, using a graph.

17. Find all real numbers x for which $|x - 2| = |x + 2|$.

18. Given that $Q(p) = p^2 - p$, find possible values of n if $Q(n) = 2$.

19. Solve $|x - 4| = 2x$.

- a. algebraically
- b. geometrically.

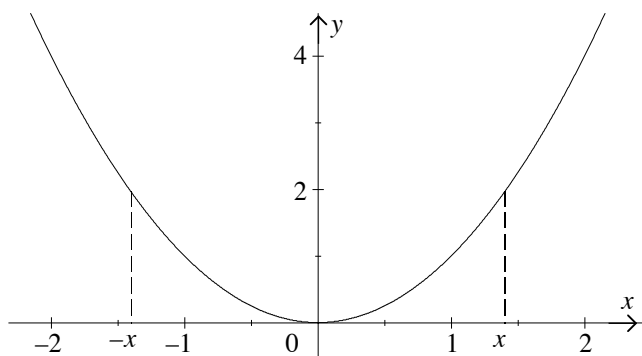
2.9 Even and odd functions

Definition:

A function, $y = f(x)$, is *even* if $f(x) = f(-x)$ for all x in the domain of f .

Geometrically, an even function is symmetrical about the y -axis (it has line symmetry).

The function $f(x) = x^2$ is an even function as $f(-x) = (-x)^2 = x^2 = f(x)$ for all values of x . We illustrate this on the following graph.



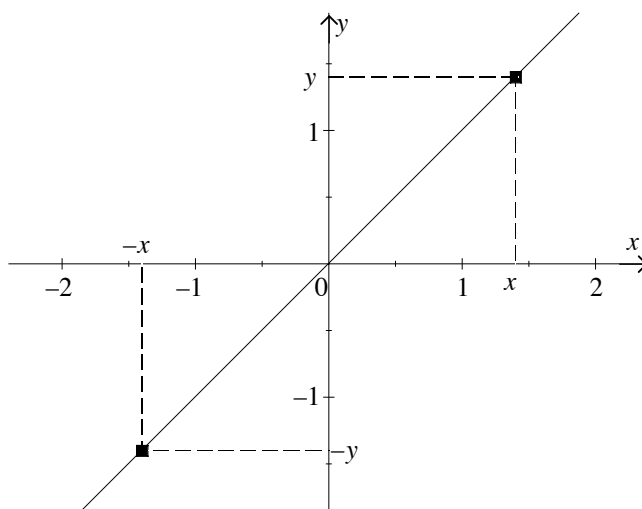
The graph of $y = x^2$.

Definition:

A function, $y = f(x)$, is *odd* if $f(-x) = -f(x)$ for all x in the domain of f .

Geometrically, an odd function is symmetrical about the origin (it has rotational symmetry).

The function $f(x) = x$ is an odd function as $f(-x) = -x = -f(x)$ for all values of x . This is illustrated on the following graph.



The graph of $y = x$.

Example

Decide whether the following functions are even, odd or neither.

1. $f(x) = 3x^2 - 4$

2. $g(x) = \frac{1}{2x}$

3. $f(x) = x^3 - x^2$.

Solution

1.

$$f(-x) = 3(-x)^2 - 4 = 3x^2 - 4 = f(x)$$

The function $f(x) = 3x^2 - 4$ is even.

2.

$$g(-x) = \frac{1}{2(-x)} = \frac{1}{-2x} = -\frac{1}{2x} = -g(x)$$

Therefore, the function g is odd.

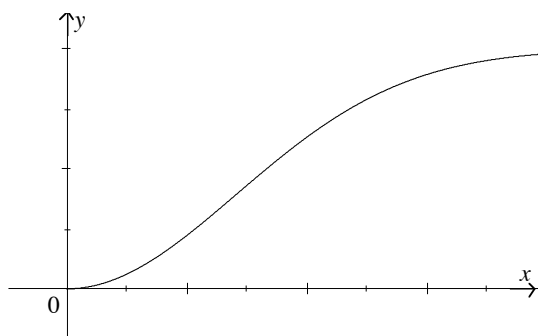
3.

$$f(-x) = (-x)^3 - (-x)^2 = -x^3 - x^2$$

This function is neither even (since $-x^3 - x^2 \neq x^3 - x^2$) nor odd (since $-x^3 - x^2 \neq -(x^3 - x^2)$).

Example

Sketched below is part of the graph of $y = f(x)$.

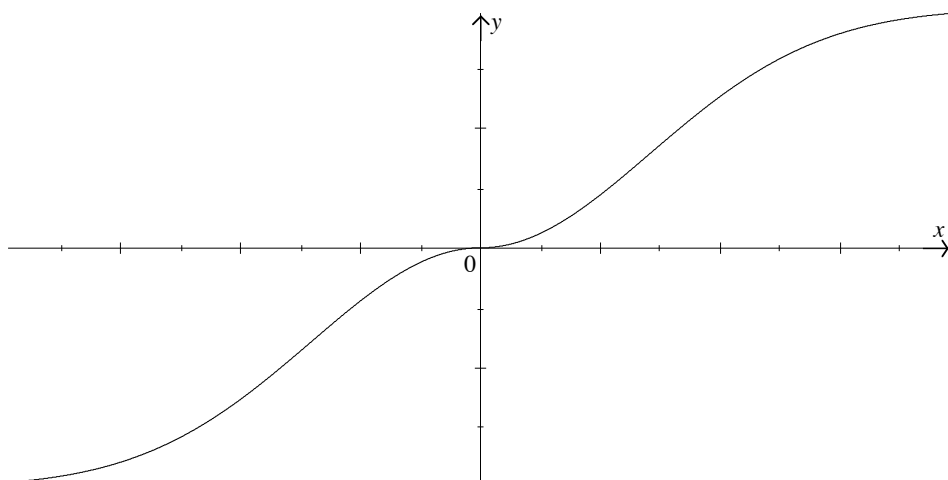


Complete the graph if $y = f(x)$ is

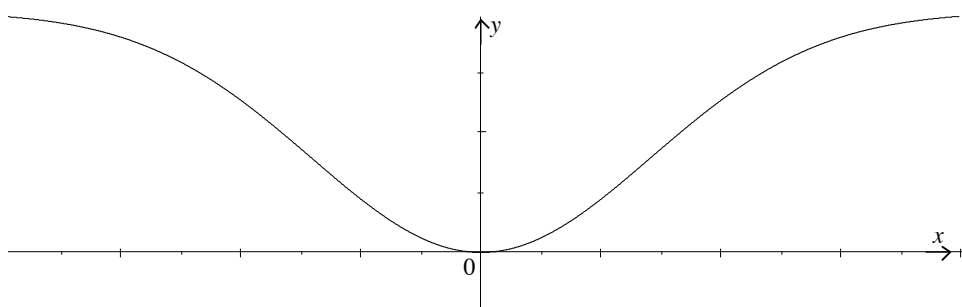
1. odd

2. even.

Solution



$y = f(x)$ is an odd function.



$y = f(x)$ is an even function.

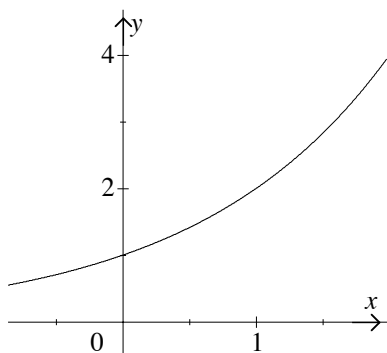
2.10 Increasing and decreasing functions

Here we will introduce the concepts of increasing and decreasing functions. In Chapter 5 we will relate these concepts to the derivative of a function.

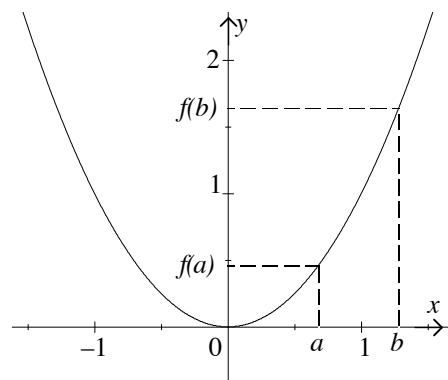
Definition:

A function is *increasing* on an interval I , if for all a and b in I such that $a < b$, $f(a) < f(b)$.

The function $y = 2^x$ is an example of a function that is increasing over its domain. The function $y = x^2$ is increasing for all real $x > 0$.



The graph of $y = 2^x$. This function is increasing for all real x .



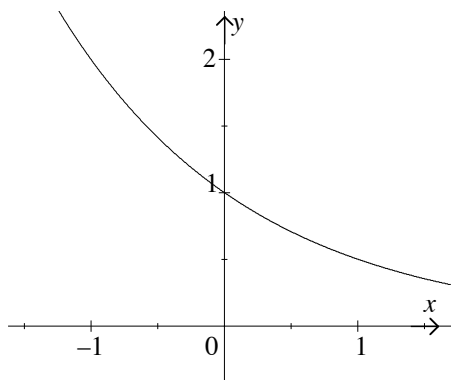
The graph of $y = x^2$. This function is increasing on the interval $x > 0$.

Notice that when a function is increasing it has a positive slope.

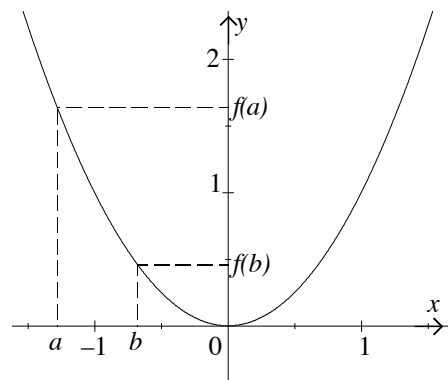
Definition:

A graph is decreasing on an interval I , if for all a and b in I such that $a < b$, $f(a) > f(b)$.

The function $y = 2^{-x}$ is decreasing over its domain. The function $y = x^2$ is decreasing on the interval $x < 0$.



The graph of $y = 2^{-x}$. This function is decreasing for all real x .



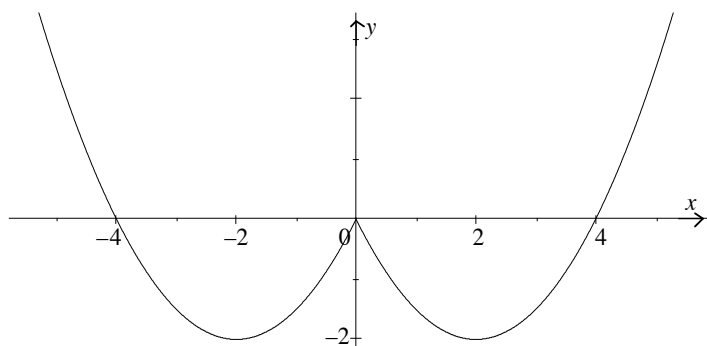
The graph of $y = x^2$. This function is decreasing on the interval $x < 0$.

Notice that if a function is decreasing then it has negative slope.

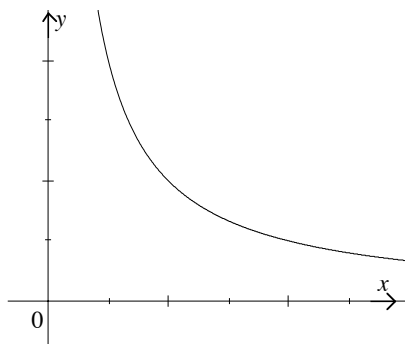
2.11 Exercises

1. Given the graph below of $y = f(x)$:
 - a. State the domain and range.
 - b. Where is the graph

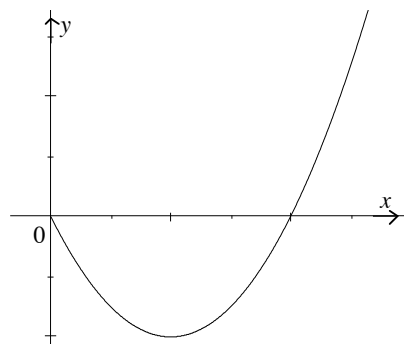
- i increasing?
- ii decreasing?
- c. if k is a constant, find the values of k such that $f(x) = k$ has
 - i no solutions
 - ii 1 solution
 - iii 2 solutions
 - iv 3 solutions
 - v 4 solutions.
- d. Is $y = f(x)$ even, odd or neither?



2. Complete the following functions if they are defined to be (a) even (b) odd.



$$y = f(x)$$



$$y = g(x)$$

3. Determine whether the following functions are odd, even or neither.
- a. $y = x^4 + 2$
 - b. $y = \sqrt{4 - x^2}$
 - c. $y = 2^x$
 - d. $y = x^3 + 3x$
 - e. $y = \frac{x}{x^2}$
 - f. $y = \frac{1}{x^2 - 4}$
 - g. $y = \frac{1}{x^2 + 4}$
 - h. $y = \frac{x}{x^3 + 3}$
 - i. $y = 2^x + 2^{-x}$
 - j. $y = |x - 1| + |x + 1|$
4. Given $y = f(x)$ is even and $y = g(x)$ is odd, prove
- a. if $h(x) = f(x) \cdot g(x)$ then $h(x)$ is odd
 - b. if $h(x) = (g(x))^2$ then $h(x)$ is even

- c. if $h(x) = \frac{f(x)}{g(x)}$, $g(x) \neq 0$, then $h(x)$ is odd
- d. if $h(x) = f(x) \cdot (g(x))^2$ then $h(x)$ is even.
5. Consider the set of all odd functions which are defined at $x = 0$. Can you prove that for every odd function in this set $f(0) = 0$? If not, give a counter-example.

3 Piecewise functions and solving inequalities

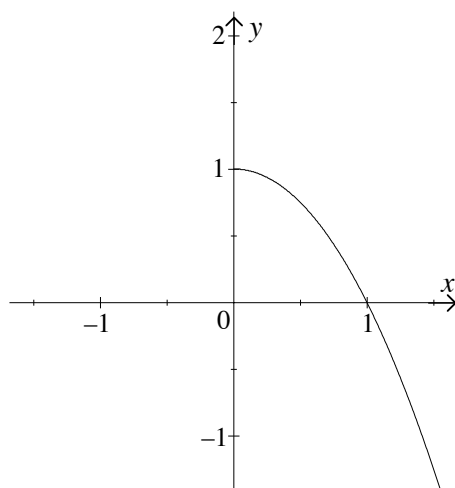
In this Chapter we will discuss functions that are defined piecewise (sometimes called piecemeal functions) and look at solving inequalities using both algebraic and graphical techniques.

3.1 Piecewise functions

3.1.1 Restricting the domain

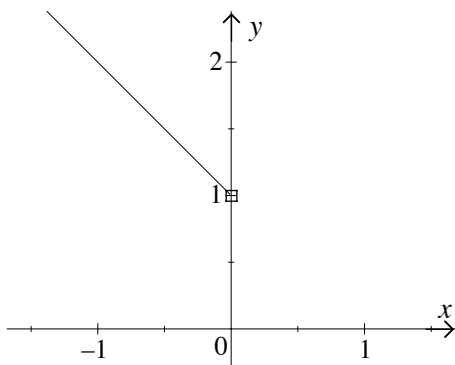
In Chapter 1 we saw how functions could be defined on a subinterval of their natural domain. This is frequently called *restricting* the domain of the function. In this Chapter we will extend this idea to define functions piecewise.

Sketch the graph of $y = 1 - x^2$ for $x \geq 0$.



The graph of $y = 1 - x^2$ for $x \geq 0$.

Sketch the graph of $y = 1 - x$ for $x < 0$.

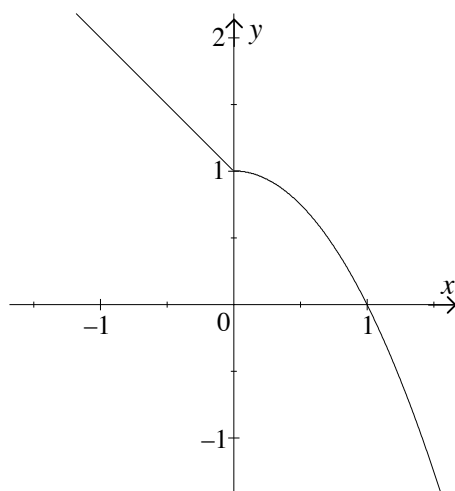


The graph of $y = 1 - x$ for $x < 0$.

We can now put these pieces together to define a function of the form

$$f(x) = \begin{cases} 1 - x^2 & \text{for } x \geq 0 \\ 1 - x & \text{for } x < 0 \end{cases}$$

We say that this function is defined *piecewise*. First note that it *is* a function; each value of x in the domain is assigned exactly one value of y . This is easy to see if we graph the function and use the vertical line test. We graph this function by graphing each piece of it in turn.



The graph shows that f defined in this way is a function. The two pieces of $y = f(x)$ meet so f is a continuous function.

The absolute value function

$$f(x) = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0 \end{cases}$$

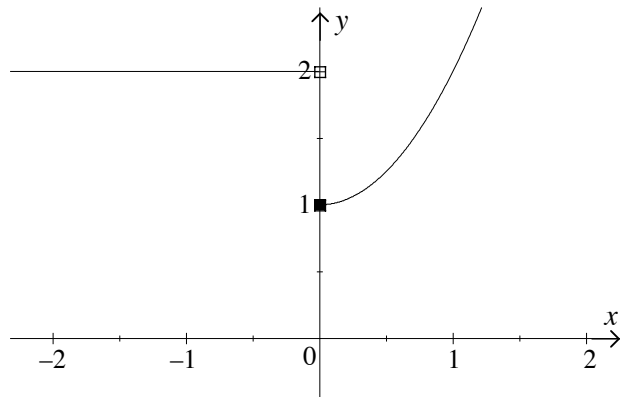
is another example of a piecewise function.

Example

Sketch the function

$$f(x) = \begin{cases} x^2 + 1 & \text{for } x \geq 0 \\ 2 & \text{for } x < 0 \end{cases}$$

Solution



This function is not continuous at $x = 0$ as the two branches of the graph do not meet.

Notice that we have put an open square (or circle) around the point $(0, 2)$ and a solid square (or circle) around the point $(0, 1)$. This is to make it absolutely clear that $f(0) = 1$ and not 2. When defining a function piecewise, we must be extremely careful to assign to each x exactly one value of y .

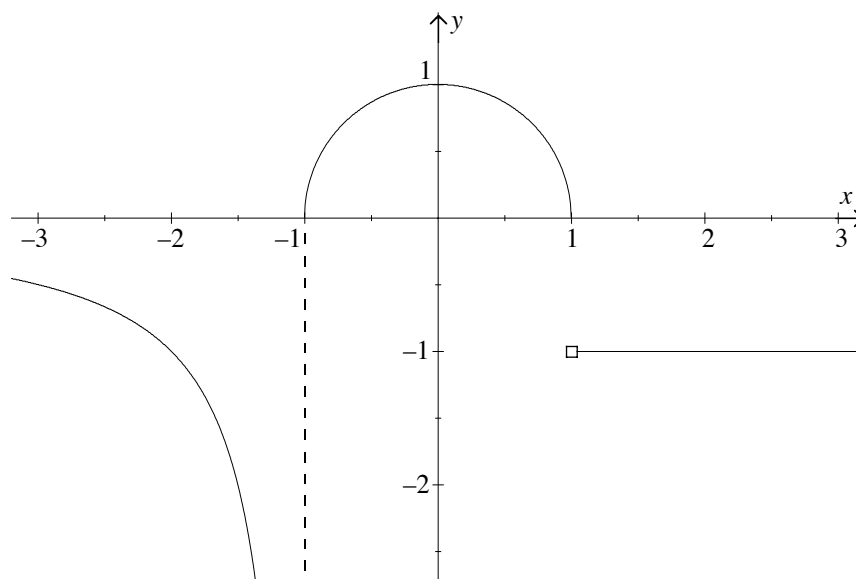
3.2 Exercises

1. For the function

$$f(x) = \begin{cases} 1 - x^2 & \text{for } x \geq 0 \\ 1 - x & \text{for } x < 0 \end{cases}$$

evaluate

- a. $2f(-1) + f(2)$
 - b. $f(a^2)$
2. For the function given in **1**, solve $f(x) = 2$.
3. Below is the graph of $y = g(x)$. Write down the rules which define $g(x)$ given that its pieces are hyperbolic, circular and linear.



4. a. Sketch the graph of $y = f(x)$ if

$$f(x) = \begin{cases} -\sqrt{4-x^2} & \text{for } -2 \leq x \leq 0 \\ x^2 - 4 & \text{for } x > 0 \end{cases}$$

- b. State the range of f .

- c. Solve

i $f(x) = 0$

ii $f(x) = -3$.

- d. Find k if $f(x) = k$ has

i 0

ii 1

iii 2 solutions.

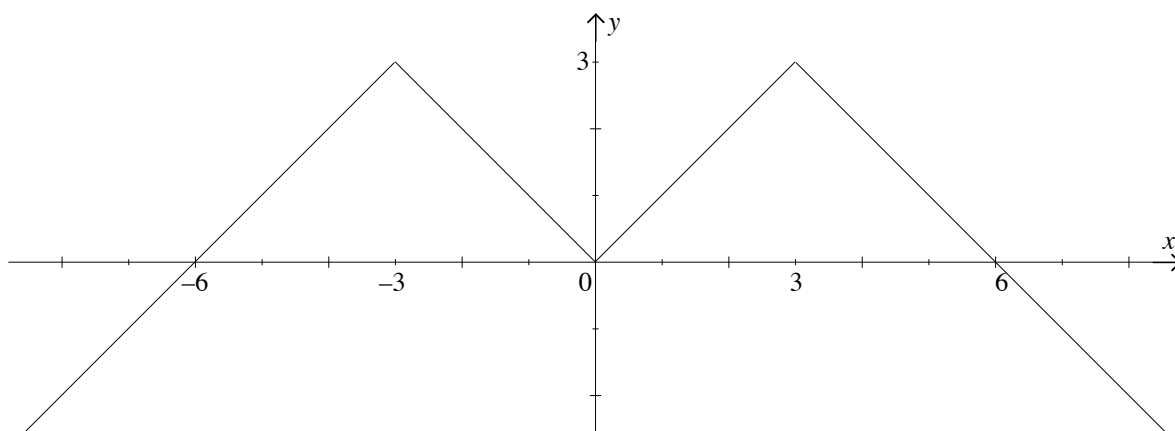
5. Sketch the graph of $y = f(x)$ if

$$f(x) = \begin{cases} 1 - |x - 1| & \text{for } x \geq 0 \\ |x + 1| & \text{for } x < 0 \end{cases}$$

6. Sketch the graph of $y = g(x)$ if

$$g(x) = \begin{cases} \frac{2}{x+2} & \text{for } x < -1 \\ 2 & \text{for } -1 \leq x < 1 \\ 2^x & \text{for } x \geq 1 \end{cases}$$

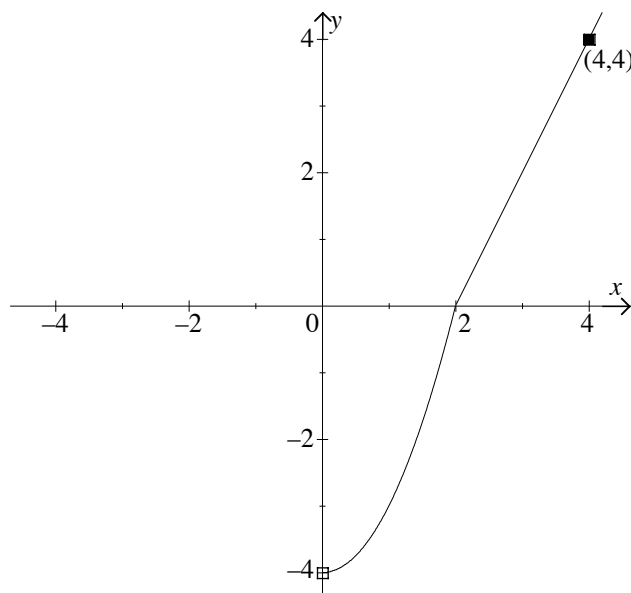
7. McMaths burgers are to modernise their logo as shown below.



Write down a piecewise function that represents this function using (a) 4 (b) 3 (c) 2 pieces (i.e. rules that define the function).

8. a. The following piecewise function is of the form

$$f(x) = \begin{cases} ax^2 + b & \text{for } 0 < x \leq 2 \\ cx + d & \text{for } x > 2 \end{cases}$$



Determine the values of a , b , c and d .

- b. Complete the graph so that $f(x)$ is an odd function defined for all real x , $x \neq 0$.
c. Write down the equations that now define $f(x)$, $x \neq 0$.

3.3 Inequalities

We can solve inequalities using both algebraic and graphical methods. Sometimes it is easier to use an algebraic method and sometimes a graphical one. For the following examples we will use both, as this allows us to make the connections between the algebra and the graphs.

Algebraic method

1. Solve $3 - 2x \geq 1$.

This is a (2 Unit) linear inequality. Remember to reverse the inequality sign when multiplying or dividing by a negative number.

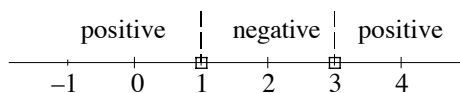
$$\begin{aligned} 3 - 2x &\geq 1 \\ -2x &\geq -2 \\ x &\leq 1 \end{aligned}$$

2. Solve $x^2 - 4x + 3 < 0$.

This is a (2 Unit) quadratic inequality. Factorise and use a number line.

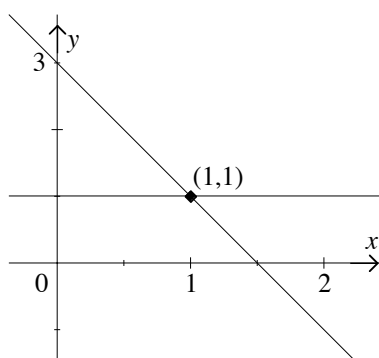
$$\begin{aligned} x^2 - 4x + 3 &< 0 \\ (x - 3)(x - 1) &< 0 \end{aligned}$$

The critical values are 1 and 3, which divide the number line into three intervals. We take points in each interval to determine the sign of the inequality; eg use $x = 0$, $x = 2$ and $x = 4$ as test values.



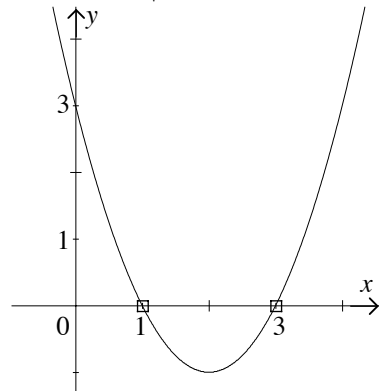
Thus, the solution is $1 < x < 3$.

Graphical method



When is the line $y = 3 - 2x$ above or on the horizontal line $y = 1$? From the graph, we see that this is true for $x \leq 1$.

Let $y = x^2 - 4x + 3$.



When does the parabola have negative y -values? OR When is the parabola under the x -axis? From the graph, we see that this happens when $1 < x < 3$.

3. Solve $\frac{1}{x-4} \leq 1$.

This is a 3 Unit inequality. There is a variable in the denominator. Remember that a denominator can never be zero, so in this case $x \neq 4$. First multiply by the square of the denominator

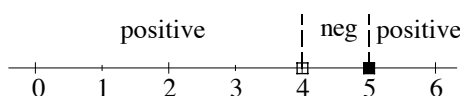
$$x - 4 \leq (x - 4)^2, x \neq 4$$

$$x - 4 \leq x^2 - 8x + 16$$

$$0 \leq x^2 - 9x + 20$$

$$0 \leq (x - 4)(x - 5)$$

Mark the critical values on the number line and test $x = 0$, $x = 4.5$ and $x = 6$.



Therefore, $x < 4$ or $x \geq 5$.

4. Solve $x - 3 < \frac{10}{x}$.

Consider $x - 3 = \frac{10}{x}$, $x \neq 0$.

Multiply by x we get

$$x^2 - 3x = 10$$

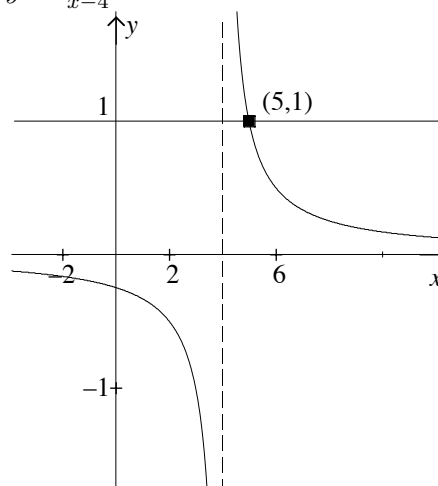
$$x^2 - 3x - 10 = 0$$

$$(x - 5)(x + 2) = 0$$

Therefore, the critical values are -2 , 0 and 5 which divide the number line into four intervals. We can use $x = -3$, $x = -1$, $x = 1$ and $x = 6$ as test values in the inequality. The points $x = -3$ and $x = 1$ satisfy the inequality, so the solution is $x < -2$ or $0 < x < 5$.

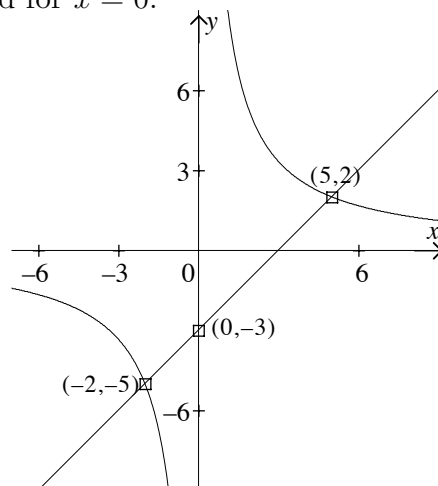
(Notice that we had to include 0 as one of our critical values.)

Let $y = \frac{1}{x-4}$.



$y = \frac{1}{x-4}$ is not defined for $x = 4$. It is a hyperbola with vertical asymptote at $x = 4$. To solve our inequality we need to find the values of x for which the hyperbola lies on or under the line $y = 1$. $(5, 1)$ is the point of intersection. So, from the graph we see that $\frac{1}{x-4} \leq 1$ when $x < 4$ or $x \geq 5$.

Sketch $y = x - 3$ and then $y = \frac{10}{x}$. Note that second of these functions is not defined for $x = 0$.



For what values of x does the line lie under the hyperbola? From the graph, we see that this happens when $x < -2$ or $0 < x < 5$.

Example

Sketch the graph of $y = |2x - 6|$.

Hence, where possible,

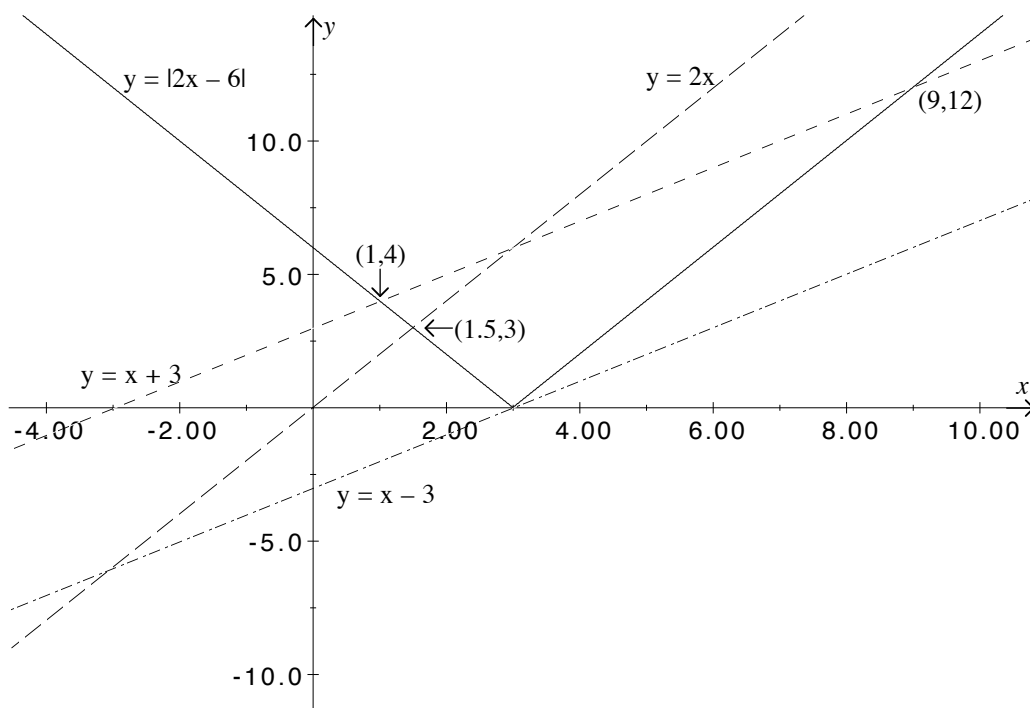
a. Solve

- i $|2x - 6| = 2x$
- ii $|2x - 6| > 2x$
- iii $|2x - 6| = x + 3$
- iv $|2x - 6| < x + 3$
- v $|2x - 6| = x - 3$

b. Determine the values of k for which $|2x - 6| = x + k$ has exactly two solutions.

Solution

$$f(x) = |2x - 6| = \begin{cases} 2x - 6 & \text{for } x \geq 3 \\ -(2x - 6) & \text{for } x < 3 \end{cases}$$



- a. i Mark in the graph of $y = 2x$. It is parallel to one arm of the absolute value graph. It has one point of intersection with $y = |2x - 6| = -2x + 6$ ($x < 3$) at $x = 1.5$.
- ii When is the absolute value graph above the line $y = 2x$? From the graph, when $x < 1.5$.

- iii $y = x + 3$ intersects $y = |2x - 6|$ twice.
To solve $|2x - 6| = x + 3$, take $|2x - 6| = 2x - 6 = x + 3$ when $x \geq 3$. This gives us the solution $x = 9$. Then take $|2x - 6| = -2x + 6 = x + 3$ when $x < 3$ which gives us the solution $x = 1$.
 - iv When is the absolute value graph below the line $y = x + 3$?
From the graph, $1 < x < 9$.
 - v $y = x - 3$ intersects the absolute value graph at $x = 3$ only.
- b. k represents the y -intercept of the line $y = x + k$. When $k = -3$, there is one point of intersection. (See (a) (v) above). For $k > -3$, lines of the form $y = x + k$ will have two points of intersection. Hence $|2x - 6| = x + k$ will have two solutions for $k > -3$.

3.4 Exercises

1. Solve
 - a. $x^2 \leq 4x$
 - b. $\frac{4p}{p+3} \leq 1$
 - c. $\frac{7}{9-x^2} > -1$
2. a. Sketch the graph of $y = 4x(x - 3)$.
b. Hence solve $4x(x - 3) \leq 0$.
3. a. Find the points of intersection of the graphs $y = 5 - x$ and $y = \frac{4}{x}$.
b. On the same set of axes, sketch the graphs of $y = 5 - x$ and $y = \frac{4}{x}$.
c. Using part (ii), or otherwise, write down all the values of x for which

$$5 - x > \frac{4}{x}$$

4. a. Sketch the graph of $y = 2^x$.
b. Solve $2^x < \frac{1}{2}$.
c. Suppose $0 < a < b$ and consider the points A($a, 2^a$) and B($b, 2^b$) on the graph of $y = 2^x$. Find the coordinates of the midpoint M of the segment AB.
Explain why

$$\frac{2^a + 2^b}{2} > 2^{\frac{a+b}{2}}$$

5. a. Sketch the graphs of $y = x$ and $y = |x - 5|$ on the same diagram.
b. Solve $|x - 5| > x$.
c. For what values of m does $mx = |x - 5|$ have exactly
 - i two solutions
 - ii no solutions
6. Solve $5x^2 - 6x - 3 \leq |8x|$.

4 Polynomials

Many of the functions we have been using so far have been polynomials. In this Chapter we will study them in more detail.

Definition

A real *polynomial*, $P(x)$, of degree n is an expression of the form

$$P(x) = p_n x^n + p_{n-1} x^{n-1} + p_{n-2} x^{n-2} + \cdots + p_2 x^2 + p_1 x + p_0$$

where $p_n \neq 0$, p_0, p_1, \dots, p_n are real and n is an integer ≥ 0 .

All polynomials are defined for all real x and are continuous functions.

We are familiar with the quadratic polynomial, $Q(x) = ax^2 + bx + c$ where $a \neq 0$. This polynomial has degree 2.

The function $f(x) = \sqrt{x} + x$ is not a polynomial as it has a power which is not an integer ≥ 0 and so does not satisfy the definition.

4.1 Graphs of polynomials and their zeros

4.1.1 Behaviour of polynomials when $|x|$ is large

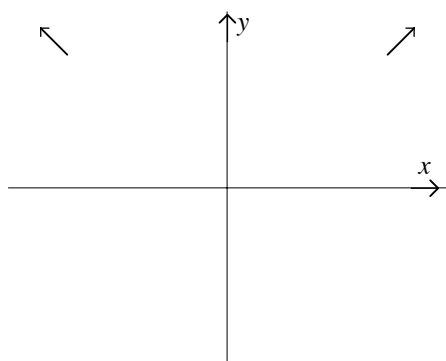
One piece of information that can be a great help when sketching a polynomial is the way it behaves for values of x when $|x|$ is large. That is, values of x which are large in magnitude.

The term of the polynomial with the highest power of x is called the *leading* or *dominant* term. For example, in the polynomial $P(x) = x^6 - 3x^4 - 1$, the term x^6 is the dominant term.

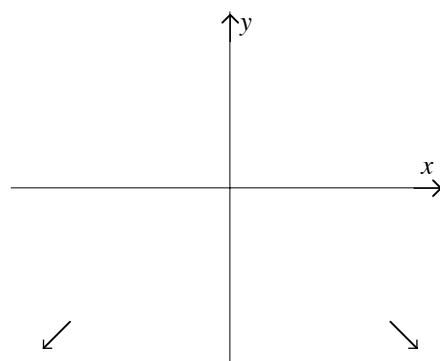
When $|x|$ is large, the dominant term determines how the graph behaves as it is so much larger in magnitude than all the other terms.

How the graph behaves for $|x|$ large depends on the *power* and *coefficient* of the dominant term.

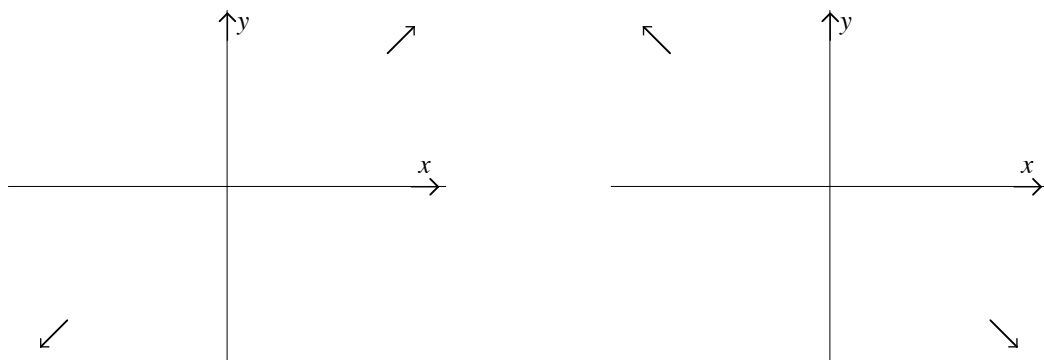
There are four possibilities which we summarise in the following diagrams:



1. Dominant term with even power and positive coefficient, eg $y = x^2$.



2. Dominant term with even power and negative coefficient, eg $Q(x) = -x^2$.



3. Dominant term with odd power and positive coefficient, eg $y = x^3$.
4. Dominant term with odd power and negative coefficient, eg $Q(x) = -x^3$.

This gives us a good start to graphing polynomials. All we need do now is work out what happens in the middle. In Chapter 5 we will use calculus methods to do this. Here we will use our knowledge of the roots of polynomials to help complete the picture.

4.1.2 Polynomial equations and their roots

If, for a polynomial $P(x)$, $P(k) = 0$ then we can say

1. $x = k$ is a root of the equation $P(x) = 0$.
2. $x = k$ is a zero of $P(x)$.
3. k is an x -intercept of the graph of $P(x)$.

4.1.3 Zeros of the quadratic polynomial

The quadratic polynomial equation $Q(x) = ax^2 + bx + c = 0$ has two roots that may be:

1. real (rational or irrational) and distinct,
2. real (rational or irrational) and equal,
3. complex (not real).

We will illustrate all of these cases with examples, and will show the relationship between the nature and number of zeros of $Q(x)$ and the x -intercepts (if any) on the graph.

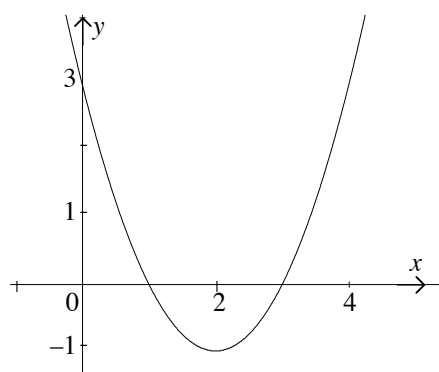
1. Let $Q(x) = x^2 - 4x + 3$.
We find the zeros of $Q(x)$ by solving the equation $Q(x) = 0$.

$$x^2 - 4x + 3 = 0$$

$$(x - 1)(x - 3) = 0$$

$$\text{Therefore } x = 1 \text{ or } 3.$$

The roots are rational (hence real) and distinct.



2. Let $Q(x) = x^2 - 4x - 3$.

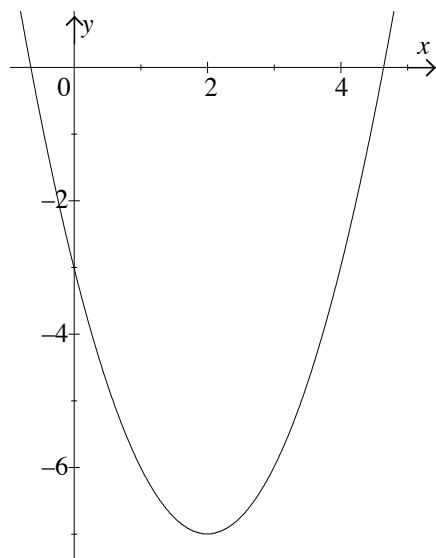
Solving the equation $Q(x) = 0$ we get,

$$x^2 - 4x - 3 = 0$$

$$x = \frac{4 \pm \sqrt{16 + 12}}{2}$$

$$\text{Therefore } x = 2 \pm \sqrt{7}.$$

The roots are irrational (hence real) and distinct.



3. Let $Q(x) = x^2 - 4x + 4$.

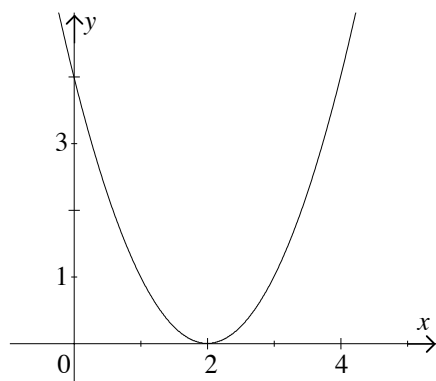
Solving the equation $Q(x) = 0$ we get,

$$x^2 - 4x + 4 = 0$$

$$(x - 2)^2 = 0$$

$$\text{Therefore } x = 2.$$

The roots are rational (hence real) and equal. $Q(x) = 0$ has a repeated or double root at $x = 2$.



Notice that the graph turns at the double root $x = 2$.

4. Let $Q(x) = x^2 - 4x + 5$.

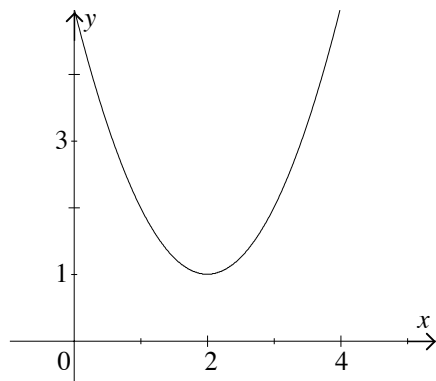
Solving the equation $Q(x) = 0$ we get,

$$x^2 - 4x + 5 = 0$$

$$x = \frac{4 \pm \sqrt{16 - 20}}{2}$$

$$\text{Therefore } x = 2 \pm \sqrt{-1}.$$

There are no real roots. In this case the roots are complex.



Notice that the graph does not intersect the x -axis. That is $Q(x) > 0$ for all real x . Therefore Q is positive definite.

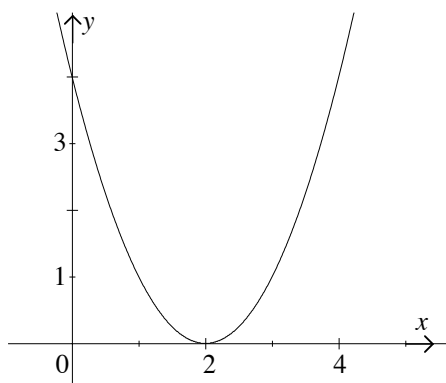
We have given above four examples of quadratic polynomials to illustrate the relationship between the zeros of the polynomials and their graphs.

In particular we saw that:

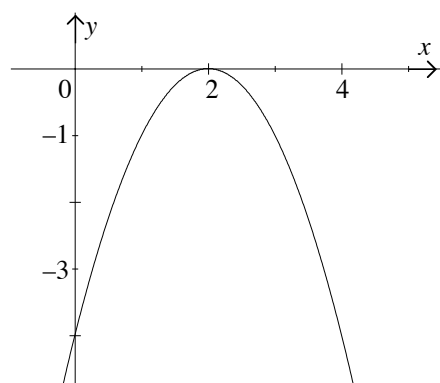
- i if the quadratic polynomial has two real distinct zeros, then the graph of the polynomial cuts the x -axis at two distinct points;
- ii if the quadratic polynomial has a real double (or repeated) zero, then the graph sits on the x -axis;
- iii if the quadratic polynomial has no real zeros, then the graph does not intersect the x -axis at all.

So far, we have only considered quadratic polynomials where the coefficient of the x^2 term is positive which gives us a graph which is *concave up*. If we consider polynomials $Q(x) = ax^2 + bx + c$ where $a < 0$ then we will have a graph which is *concave down*.

For example, the graph of $Q(x) = -(x^2 - 4x + 4)$ is the reflection in the x -axis of the graph of $Q(x) = x^2 - 4x + 4$. (See Chapter 2.)



The graph of $Q(x) = x^2 - 4x + 4$.



The graph of $Q(x) = -(x^2 - 4x + 4)$.

4.1.4 Zeros of cubic polynomials

A real cubic polynomial has an equation of the form

$$P(x) = ax^3 + bx^2 + cx + d$$

where $a \neq 0$, a , b , c and d are real. It has 3 zeros which may be:

- i 3 real distinct zeros;
- ii 3 real zeros, all of which are equal (3 equal zeros);
- iii 3 real zeros, 2 of which are equal;
- iv 1 real zero and 2 complex zeros.

We will illustrate these cases with the following examples:

1. Let $Q(x) = 3x^3 - 3x$.

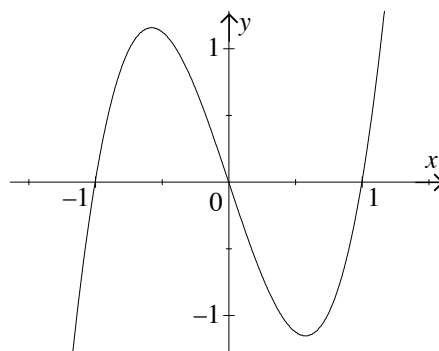
Solving the equation $Q(x) = 0$ we get:.

$$3x^3 - 3x = 0$$

$$3x(x-1)(x+1) = 0$$

Therefore $x = -1$ or 0 or 1

The roots are real (in fact rational) and distinct.

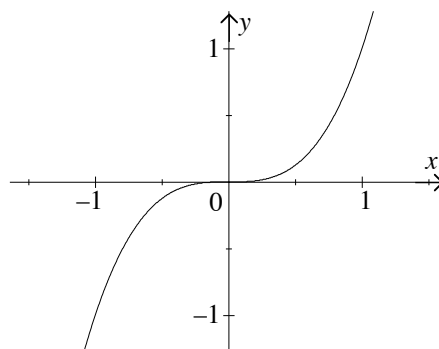


2. Let $Q(x) = x^3$.

Solving $Q(x) = 0$ we get that $x^3 = 0$.

We can write this as $(x-0)^3 = 0$.

So, this equation has three equal real roots at $x = 0$.



3. Let $Q(x) = x^3 - x^2$.

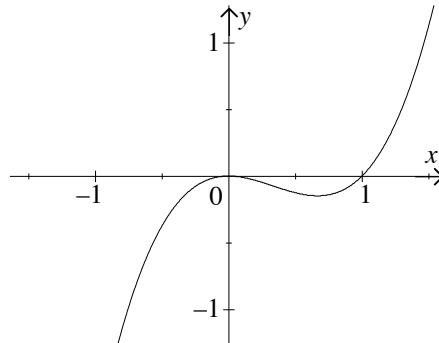
Solving the equation $Q(x) = 0$ we get,

$$x^3 - x^2 = 0$$

$$x^2(x-1) = 0$$

Therefore $x = 0$ or 1 .

The roots are real with a double root at $x = 0$ and a single root at $x = 1$.



The graph turns at the double root.

4. Let $Q(x) = x^3 + x$.

Solving the equation $Q(x) = 0$ we get,

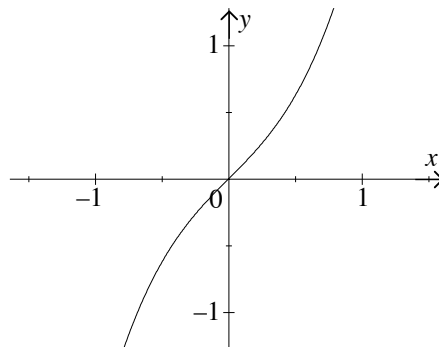
$$x^3 + x = 0$$

$$x(x^2 + 1) = 0$$

Therefore $x = 0$.

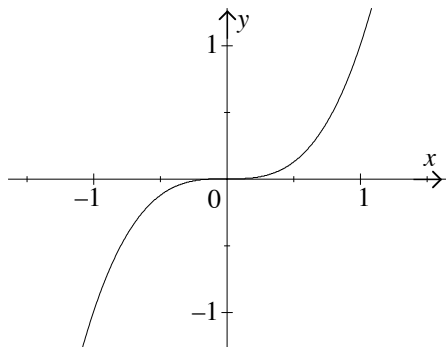
There is one real root at $x = 0$.

$x^2 + 1 = 0$ does not have any real solutions.

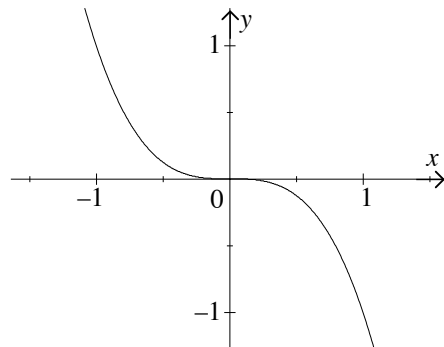


The graph intersects the x -axis once only.

Again, in the above examples we have looked only at cubic polynomials where the coefficient of the x^3 term is positive. If we consider the polynomial $P(x) = -x^3$ then the graph of this polynomial is the reflection of the graph of $P(x) = x^3$ in the x -axis.



The graph of $Q(x) = x^3$.



The graph of $Q(x) = -x^3$.

4.2 Polynomials of higher degree

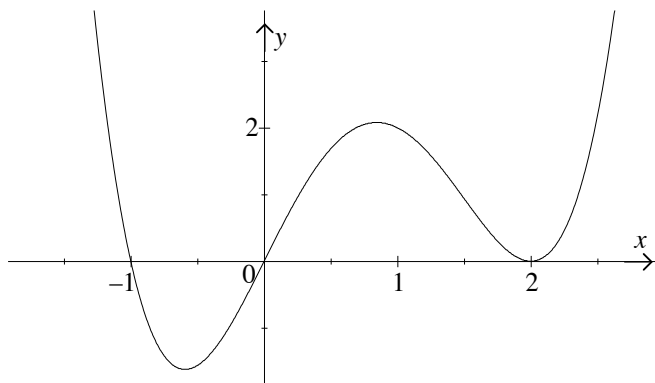
We will write down a few rules that we can use when we have a polynomial of degree ≥ 3 .

If $P(x)$ is a real polynomial of degree n then:

1. $P(x) = 0$ has at most n real roots;
2. if $P(x) = 0$ has a repeated root with an even power then the graph of $P(x)$ turns at this repeated root;
3. if $P(x) = 0$ has a repeated root with an odd power then the graph of $P(x)$ has a horizontal point of inflection at this repeated root.

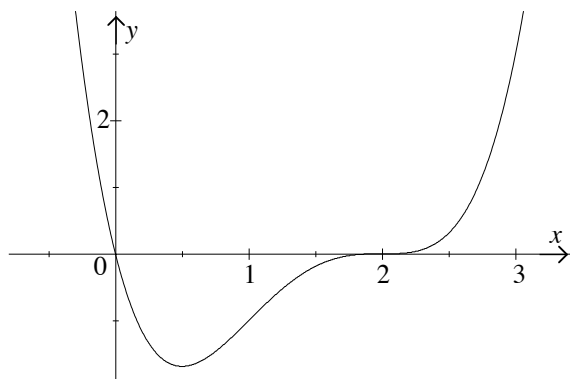
For example, **1.** tells us that if we have a quartic polynomial equation $f(x) = 0$. Then we know that $f(x) = 0$ has ≤ 4 real roots.

We can illustrate **2.** by the sketching $f(x) = x(x - 2)^2(x + 1)$. Notice how the graph sits on the x -axis at $x = 2$.



The graph of $f(x) = x(x + 1)(x - 2)^2$.

We illustrate **3.** by sketching the graph of $f(x) = x(x - 2)^3$. Notice the horizontal point of inflection at $x = 2$.



The graph of $f(x) = x(x - 2)^3$.

4.3 Exercises

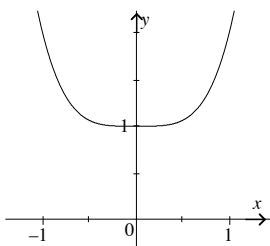
1. Sketch the graphs of the following polynomials if $y = P(x)$ is:

- a. $x(x + 1)(x - 3)$
- b. $x(x + 1)(3 - x)$
- c. $(x + 1)^2(x - 3)$
- d. $(x + 1)(x^2 - 4x + 5)$

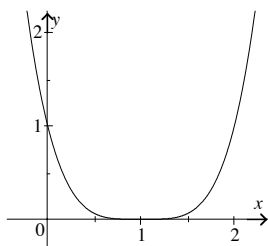
2. The graphs of the following quartic polynomials are sketched below. Match the graph with the polynomial.

- a. $y = x^4$ b. $y = x^4 - 1$ c. $y = x^4 + 1$ d. $y = 1 - x^4$ e. $y = (x - 1)^4$ f. $y = (x + 1)^4$

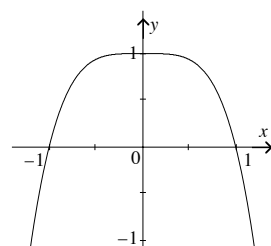
i



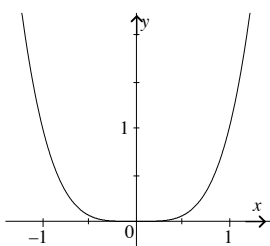
ii



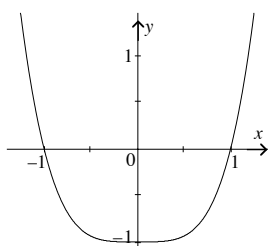
iii



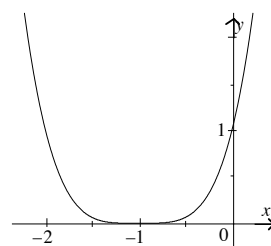
iv



v

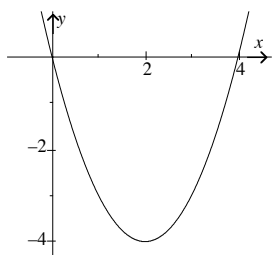


vi

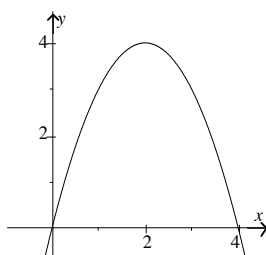


3. Sketch the graphs of the following quartic polynomials if $y = C(x)$ is:
- $x(x-1)(x+2)(x+3)$
 - $x(x-1)(x+2)(3-x)$
 - $x^2(x-1)(x-3)$
 - $(x+1)^2(x-3)^2$
 - $(x+1)^3(x-3)$
 - $(x+1)^3(3-x)$
 - $x(x+1)(x^2-4x+5)$
 - $x^2(x^2-4x+5)$.
4. By sketching the appropriate polynomial, solve:
- $x^2 - 4x - 12 < 0$
 - $(x+2)(x-3)(5-x) > 0$
 - $(x+2)^2(5-x) > 0$
 - $(x+2)^3(5-x) \geq 0$.
5. For what values of k will $P(x) \geq 0$ for all real x if $P(x) = x^2 - 4x - 12 + k$?
6. The diagrams show the graph of $y = P(x)$ where $P(x) = a(x-b)(x-c)^d$. In each case determine possible values for a , b , c and d .

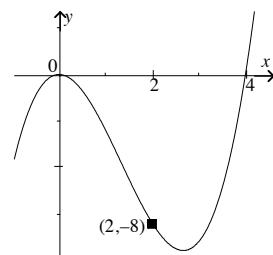
a.



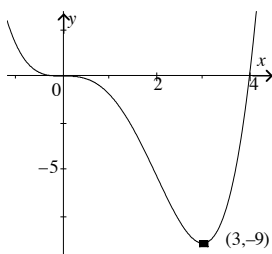
b.



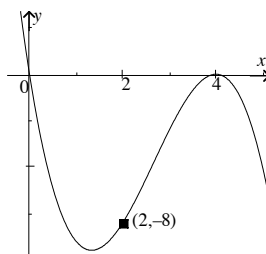
c.



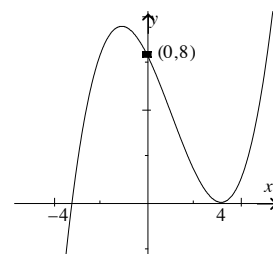
d.



e.

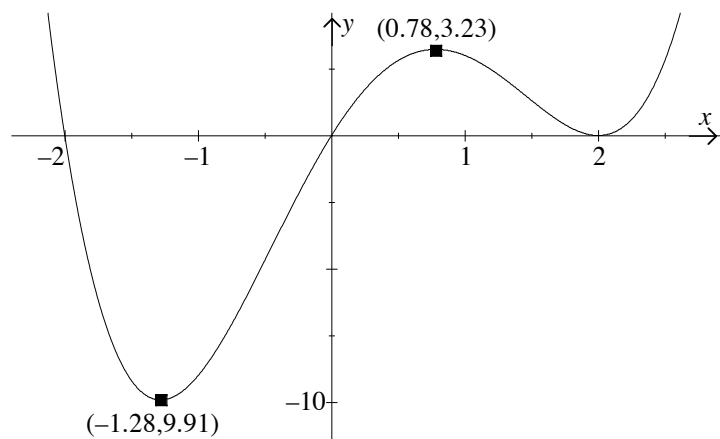


f.



7. The graph of the polynomial $y = f(x)$ is given below. It has a local maximum and minimum as marked. Use the graph to answer the following questions.
- State the roots of $f(x) = 0$.
 - What is the value of the repeated root.
 - For what values of k does the equation $f(x) = k$ have exactly 3 solutions.

- d. Solve the inequality $f(x) < 0$.
- e. What is the *least* possible degree of $f(x)$?
- f. State the value of the constant of $f(x)$.
- g. For what values of k is $f(x) + k \geq 0$ for all real x .



The graph of the polynomial $y = f(x)$

4.4 Factorising polynomials

So far for the most part, we have looked at polynomials which were already factorised. In this section we will look at methods which will help us factorise polynomials with degree > 2 .

4.4.1 Dividing polynomials

Suppose we have two polynomials $P(x)$ and $A(x)$, with the degree of $P(x) \geq$ the degree of $A(x)$, and $P(x)$ is divided by $A(x)$. Then

$$\frac{P(x)}{A(x)} = Q(x) + \frac{R(x)}{A(x)},$$

where $Q(x)$ is a polynomial called the *quotient* and $R(x)$ is a polynomial called the *remainder*, with the degree of $R(x) < \text{degree of } A(x)$.

We can rewrite this as

$$P(x) = A(x) \cdot Q(x) + R(x).$$

For example: If $P(x) = 2x^3 + 4x + 3$ and $A(x) = x - 2$, then $P(x)$ can be divided by $A(x)$ as follows:

$$\begin{array}{r}
 2x^2 + 4x + 12 \\
 x - 2 \overline{) 2x^3 + 0x^2 + 4x - 3} \\
 \underline{2x^3 - 4x^2} \\
 4x^2 + 4x - 3 \\
 \underline{4x^2 - 8x} \\
 12x - 3 \\
 \underline{12x - 24} \\
 21
 \end{array}$$

The quotient is $2x^2 + 4x + 12$ and the remainder is 21. We have

$$\frac{2x^3 + 4x + 3}{x - 2} = 2x^2 + 4x + 12 + \frac{21}{x - 2}.$$

This can be written as

$$2x^3 + 4x - 3 = (x - 2)(2x^2 + 4x + 12) + 21.$$

Note that the degree of the "polynomial" 21 is 0.

4.4.2 The Remainder Theorem

If the polynomial $f(x)$ is divided by $(x - a)$ then the remainder is $f(a)$.

Proof:

Following the above, we can write

$$f(x) = A(x) \cdot Q(x) + R(x),$$

where $A(x) = (x - a)$. Since the degree of $A(x)$ is 1, the degree of $R(x)$ is zero. That is, $R(x) = r$ where r is a constant.

$$\begin{aligned} f(x) &= (x - a)Q(x) + r \quad \text{where } r \text{ is a constant.} \\ f(a) &= 0 \cdot Q(a) + r \\ &= r \end{aligned}$$

So, if $f(x)$ is divided by $(x - a)$ then the remainder is $f(a)$.

Example

Find the remainder when $P(x) = 3x^4 - x^3 + 30x - 1$ is divided by **a.** $x + 1$, **b.** $2x - 1$.

Solution

a. Using the Remainder Theorem:

$$\begin{aligned} \text{Remainder} &= P(-1) \\ &= 3 - (-1) - 30 - 1 \\ &= -27 \end{aligned}$$

b.

$$\begin{aligned} \text{Remainder} &= P\left(\frac{1}{2}\right) \\ &= 3\left(\frac{1}{2}\right)^4 - \left(\frac{1}{2}\right)^3 + 30\left(\frac{1}{2}\right) - 1 \\ &= \frac{3}{16} - \frac{1}{8} + 15 - 1 \\ &= 14\frac{1}{16} \end{aligned}$$

Example

When the polynomial $f(x)$ is divided by $x^2 - 4$, the remainder is $5x + 6$. What is the remainder when $f(x)$ is divided by $(x - 2)$?

Solution

Write $f(x) = (x^2 - 4) \cdot q(x) + (5x + 6)$. Then

$$\begin{aligned}\text{Remainder} &= f(2) \\ &= 0 \cdot q(2) + 16 \\ &= 16\end{aligned}$$

A consequence of the Remainder Theorem is the Factor Theorem which we state below.

4.4.3 The Factor Theorem

If $x = a$ is a zero of $f(x)$, that is $f(a) = 0$, then $(x - a)$ is a factor of $f(x)$ and $f(x)$ may be written as

$$f(x) = (x - a)q(x)$$

for some polynomial $q(x)$.

Also, if $(x - a)$ and $(x - b)$ are factors of $f(x)$ then $(x - a)(x - b)$ is a factor of $f(x)$ and

$$f(x) = (x - a)(x - b) \cdot Q(x)$$

for some polynomial $Q(x)$.

Another useful fact about zeros of polynomials is given below for a polynomial of degree 3.

If a (real) polynomial

$$P(x) = ax^3 + bx^2 + cx + d,$$

where $a \neq 0$, a , b , c and d are real, has exactly 3 real zeros α , β and γ , then

$$P(x) = a(x - \alpha)(x - \beta)(x - \gamma) \quad (1)$$

Furthermore, by expanding the right hand side of (1) and equating coefficients we get:

i

$$\alpha + \beta + \gamma = -\frac{b}{a};$$

ii

$$\alpha\beta + \alpha\gamma + \beta\gamma = \frac{c}{a};$$

iii

$$\alpha\beta\gamma = -\frac{d}{a}.$$

This result can be extended for polynomials of degree n . We will give the partial result for $n = 4$.

If

$$P(x) = ax^4 + bx^3 + cx^2 + dx + e$$

is a polynomial of degree 4 with real coefficients, and $P(x)$ has four real zeros α , β , γ and δ , then

$$P(x) = a(x - \alpha)(x - \beta)(x - \gamma)(x - \delta)$$

and expanding and equating as above gives

$$\alpha\beta\gamma\delta = \frac{e}{a}.$$

If $a = 1$ and the equation $P(x) = 0$ has a root which is an integer, then that integer must be a factor of the constant term. This gives us a place to start when looking for factors of a polynomial. That is, we look at all the factors of the constant term to see which ones (if any) are roots of the equation $P(x) = 0$.

Example

Let $f(x) = 4x^3 - 8x^2 - x + 2$

- a. Factorise $f(x)$.
- b. Sketch the graph of $y = f(x)$.
- c. Solve $f(x) \geq 0$.

Solution

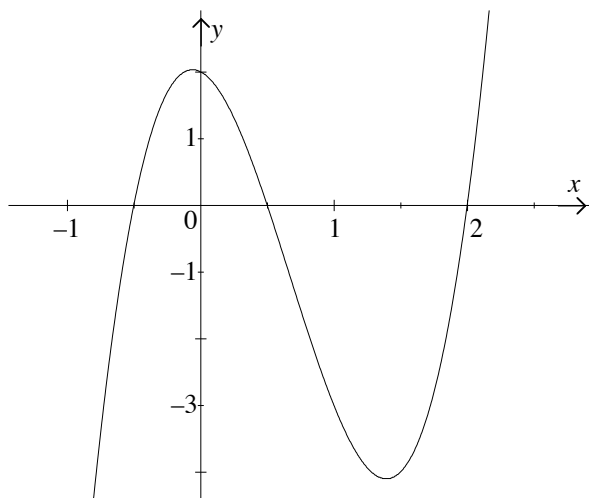
- a. Consider the factors of the constant term, 2. We check to see if ± 1 and ± 2 are solutions of the equation $f(x) = 0$ by substitution. Since $f(2) = 0$, we know that $(x - 2)$ is a factor of $f(x)$. We use long division to determine the quotient.

$$\begin{array}{r} 4x^2 - 1 \\ x - 2 \overline{) 4x^3 - 8x^2 - x + 2} \\ \underline{4x^3 - 8x^2} \\ -x + 2 \\ \underline{-x + 2} \\ 0 \end{array}$$

So,

$$\begin{aligned} f(x) &= (x - 2)(4x^2 - 1) \\ &= (x - 2)(2x - 1)(2x + 1) \end{aligned}$$

b.



The graph of $f(x) = 4x^3 - 8x^2 - x + 2$.

c. $f(x) \geq 0$ when $-\frac{1}{2} \leq x \leq \frac{1}{2}$ or $x \geq 2$.

Example

Show that $(x - 2)$ and $(x - 3)$ are factors of $P(x) = x^3 - 19x + 30$, and hence solve $x^3 - 19x + 30 = 0$.

Solution

$P(2) = 8 - 38 + 30 = 0$ and $P(3) = 27 - 57 + 30 = 0$ so $(x - 2)$ and $(x - 3)$ are both factors of $P(x)$ and $(x - 2)(x - 3) = x^2 - 5x + 6$ is also a factor of $P(x)$. Long division of $P(x)$ by $x^2 - 5x + 6$ gives a quotient of $(x + 5)$.

So,

$$P(x) = x^3 - 19x + 30 = (x - 2)(x - 3)(x + 5).$$

Solving $P(x) = 0$ we get $(x - 2)(x - 3)(x + 5) = 0$.

That is, $x = 2$ or $x = 3$ or $x = -5$.

Instead of using long division we could have used the facts that

- i the polynomial cannot have more than three real zeros;
- ii the product of the zeros must be equal to -30 .

Let α be the unknown root.

Then $2 \cdot 3 \cdot \alpha = -30$, so that $\alpha = -5$. Therefore the solution of $P(x) = x^3 - 19x + 30 = 0$ is $x = 2$ or $x = 3$ or $x = -5$.

4.5 Exercises

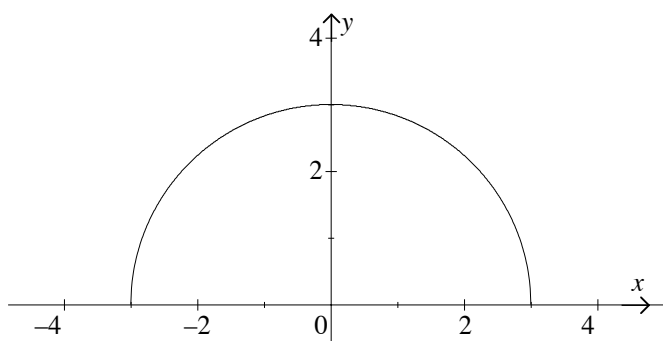
1. When the polynomial $P(x)$ is divided by $(x - a)(x - b)$ the quotient is $Q(x)$ and the remainder is $R(x)$.
 - a. Explain why $R(x)$ is of the form $mx + c$ where m and c are constants.
 - b. When a polynomial is divided by $(x - 2)$ and $(x - 3)$, the remainders are 4 and 9 respectively. Find the remainder when the polynomial is divided by $x^2 - 5x + 6$.
 - c. When $P(x)$ is divided by $(x - a)$ the remainder is a^2 . Also, $P(b) = b^2$. Find $R(x)$ when $P(x)$ is divided by $(x - a)(x - b)$.
2.
 - a. Divide the polynomial $f(x) = 2x^4 + 13x^3 + 18x^2 + x - 4$ by $g(x) = x^2 + 5x + 2$. Hence write $f(x) = g(x)q(x) + r(x)$ where $q(x)$ and $r(x)$ are polynomials.
 - b. Show that $f(x)$ and $g(x)$ have no common zeros. (Hint: Assume that α is a common zero and show by contradiction that α does not exist.)
3. For the following polynomials,
 - i factorise
 - ii solve $P(x) = 0$
 - iii sketch the graph of $y = P(x)$.
 - a. $P(x) = x^3 - x^2 - 10x - 8$
 - b. $P(x) = x^3 - x^2 - 16x - 20$
 - c. $P(x) = x^3 + 4x^2 - 8$
 - d. $P(x) = x^3 - x^2 + x - 6$
 - e. $P(x) = 2x^3 - 3x^2 - 11x + 6$

5 Solutions to exercises

1.4 Solutions

1. a. The domain of $f(x) = \sqrt{9 - x^2}$ is all real x where $-3 \leq x \leq 3$. The range is all real y such that $0 \leq y \leq 3$.

b.

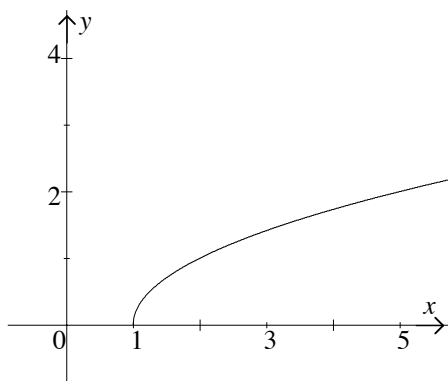


The graph of $f(x) = \sqrt{9 - x^2}$.

2.

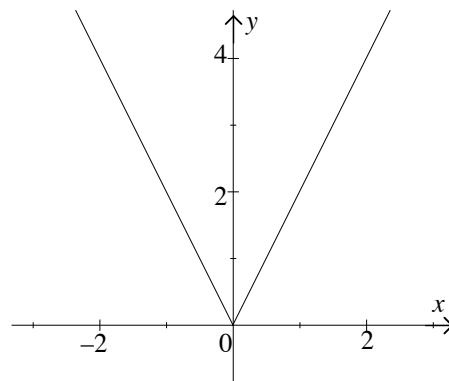
$$\begin{aligned} \frac{\psi(x+h) - \psi(x)}{h} &= \frac{(x+h)^2 + 5 - (x^2 + 5)}{h} \\ &= \frac{x^2 + 2xh + h^2 + 5 - x^2 - 5}{h} \\ &= \frac{h^2 + 2xh}{h} \\ &= h + 2x \end{aligned}$$

3. a.



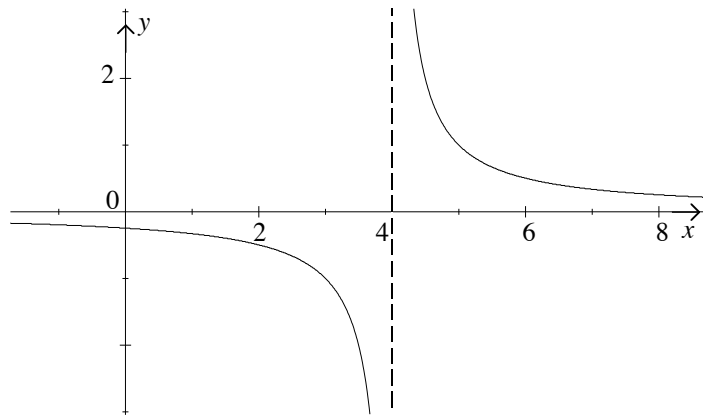
The graph of $y = \sqrt{x - 1}$. The domain is all real $x \geq 1$ and the range is all real $y \geq 0$.

b.



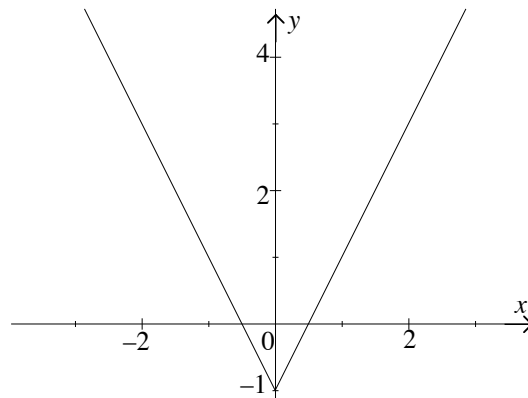
The graph of $y = |2x|$. Its domain is all real x and range all real $y \geq 0$.

c.



The graph of $y = \frac{1}{x-4}$. The domain is all real $x \neq 4$ and the range is all real $y \neq 0$.

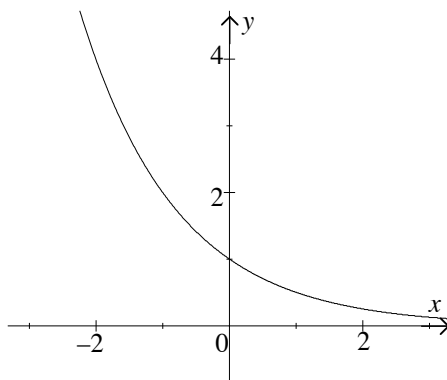
d.



The graph of $y = |2x| - 1$. The domain is all real x , and the range is all real $y \geq -1$.

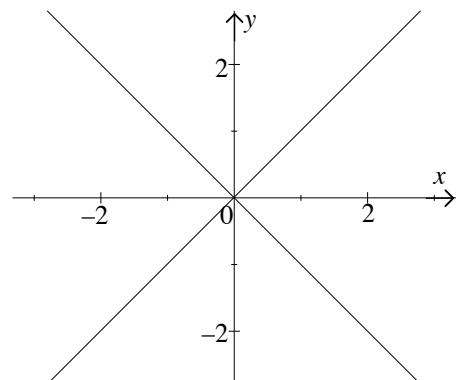
4. a. The perpendicular distance d from $(0, 0)$ to $x + y + k = 0$ is $d = \left| \frac{k}{\sqrt{2}} \right|$.
- b. For the line $x + y + k = 0$ to cut the circle in two distinct points $d < 2$. ie $|k| < 2\sqrt{2}$ or $-2\sqrt{2} < k < 2\sqrt{2}$.

5. a.



The graph of $y = \left(\frac{1}{2}\right)^x$.

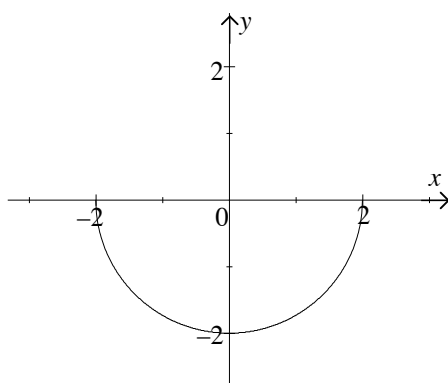
b.



The graph of $y^2 = x^2$.

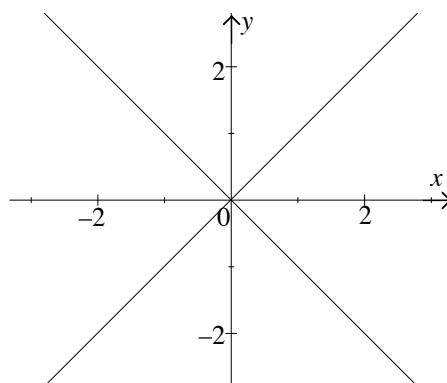
6. $y^2 = x^3$ is not a function.

7. a.



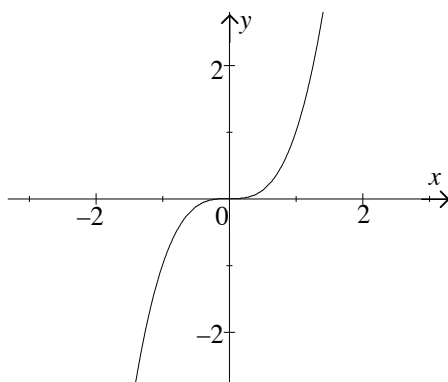
The graph of $y = -\sqrt{4 - x^2}$. This is a function with the domain: all real x such that $-2 \leq x \leq 2$ and range: all real y such that $-2 \leq y \leq 0$.

b.



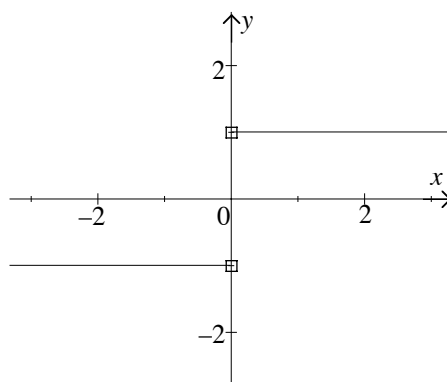
The graph of $|x| - |y| = 0$. This is not the graph of a function.

c.



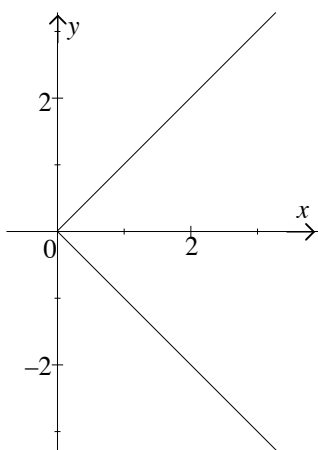
The graph of $y = x^3$. This is a function with the domain: all real x and range: all real y .

d.



The graph of $y = \frac{x}{|x|}$. This is the graph of a function which is not defined at $x = 0$. Its domain is all real $x \neq 0$, and range is $y = \pm 1$.

e.



The graph of $|y| = x$. This is not the graph of a function.

8.

$$\begin{aligned}
 A\left(\frac{1}{p}\right) &= \left(\frac{1}{p}\right)^2 + 2 + \frac{1}{\left(\frac{1}{p}\right)^2} \\
 &= \frac{1}{p^2} + 2 + \frac{1}{\frac{1}{p^2}} \\
 &= \frac{1}{p^2} + 2 + p^2 \\
 &= A(p)
 \end{aligned}$$

9. a. The values of x in the interval $0 < x < 4$ are not in the domain of the function.

b. $x = 1$ and $x = -1$ are not in the domain of the function.

10. a. $\phi(3) + \phi(4) + \phi(5) = \log(2.5)$

b. $\phi(3) + \phi(4) + \phi(5) + \cdots + \phi(n) = \log\left(\frac{n}{2}\right)$

11. a. $y = 3$ when $z = 3$.

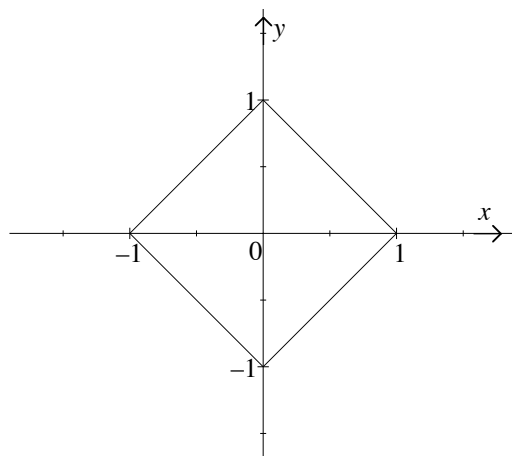
b. i $L(M(x)) = 2(x^2 - x) + 1$

ii $M(L(x)) = 4x^2 + 2x$

12. a. $a = 2$, $b = 2$ so the equations is $y = 2x^2 - 2$.

b. $a = 5$, $b = 1$ so the equation is $y = \frac{5}{x^2+1}$.

13. b.



The graph of $|x| + |y| = 1$.

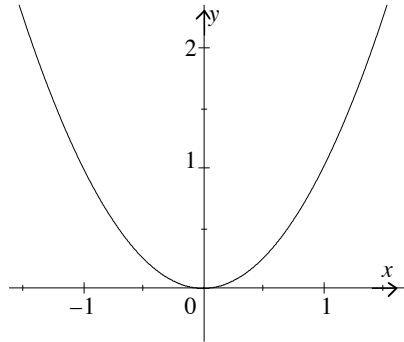
14. $S(n-1) = \frac{n-1}{2n-1}$

Hence

$$\begin{aligned}
 S(n) - S(n-1) &= \frac{n}{2n+1} - \frac{n-1}{2n-1} \\
 &= \frac{n(2n-1) - (2n+1)(n-1)}{(2n-1)(2n+1)} \\
 &= \frac{2n^2 - n - (2n^2 - n - 1)}{(2n-1)(2n+1)} \\
 &= \frac{1}{(2n-1)(2n+1)}
 \end{aligned}$$

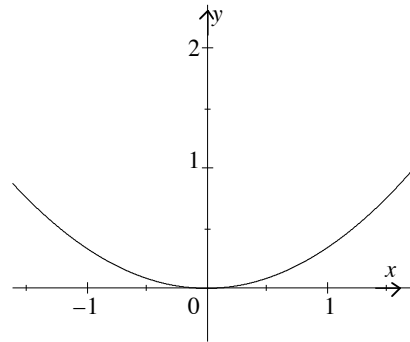
2.8 Solutions

1. a.



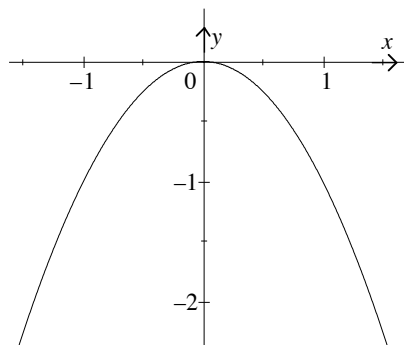
The graph of $y = x^2$.

b.



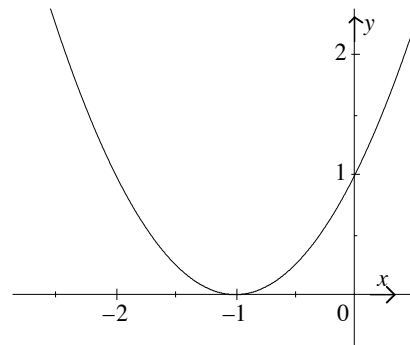
The graph of $y = \frac{x^2}{3}$.

c.



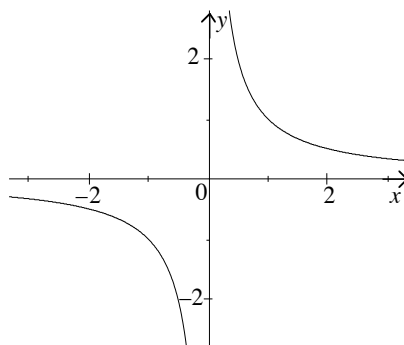
The graph of $y = -x^2$.

d.



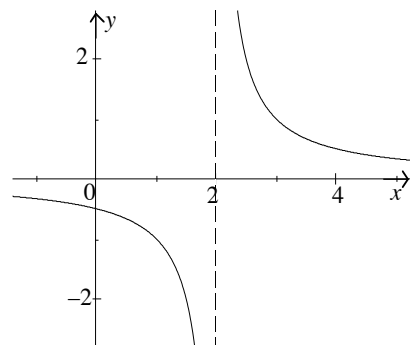
The graph of $y = (x + 1)^2$.

2. a.



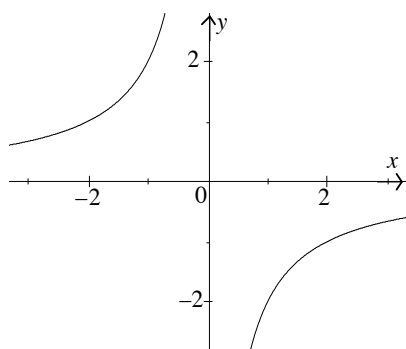
The graph of $y = \frac{1}{x}$.

b.



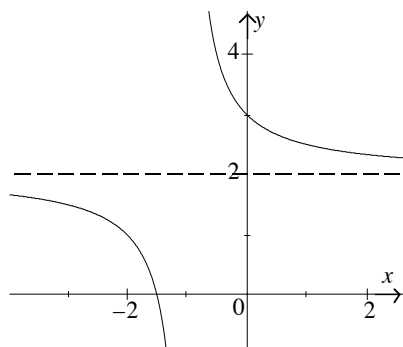
The graph of $y = \frac{1}{x-2}$.

c.



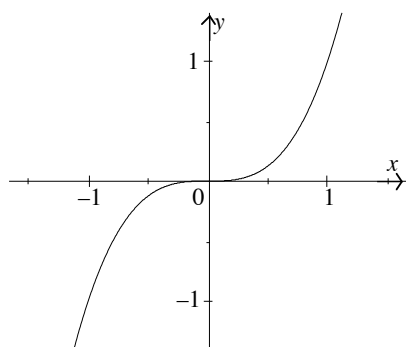
The graph of $y = \frac{-2}{x}$.

d.



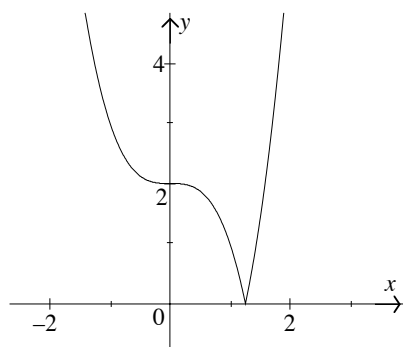
The graph of $y = \frac{1}{x+1} + 2$.

3. a.



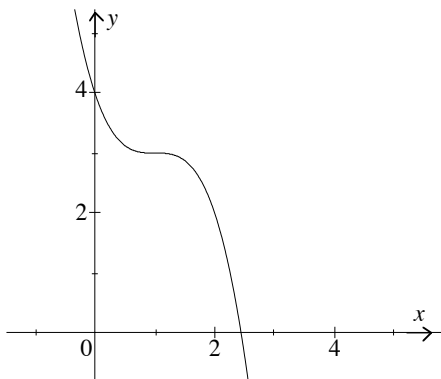
The graph of $y = x^3$.

b.



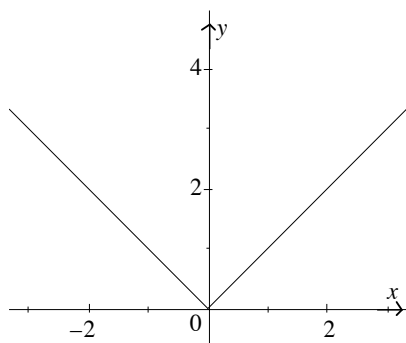
The graph of $y = |x^3 - 2|$.

c.



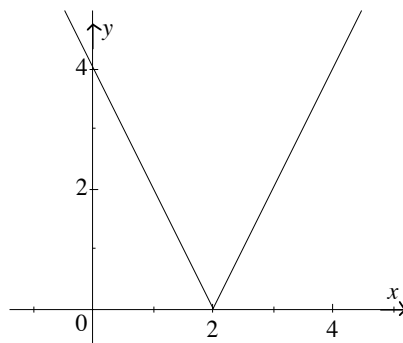
The graph of $y = 3 - (x - 1)^3$.

4. a.



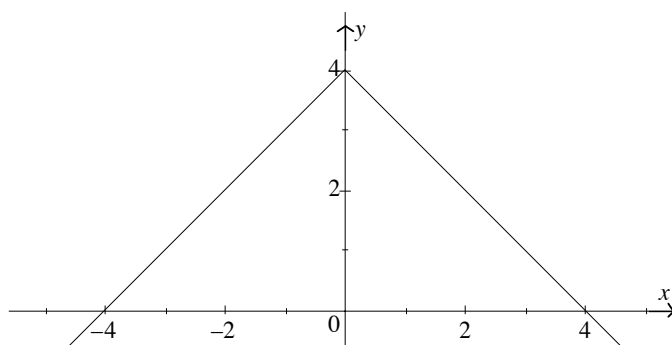
The graph of $y = |x|$.

b.



The graph of $y = 2|x - 2|$.

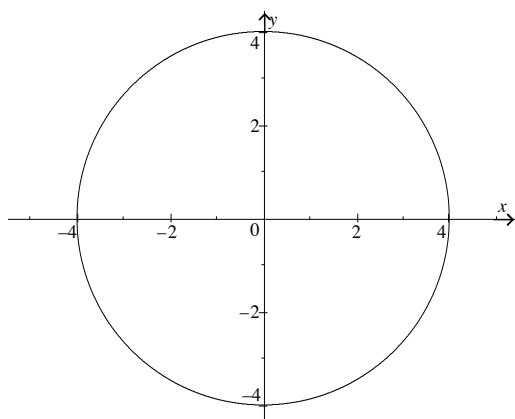
c.



The graph of $y = 4 - |x|$.

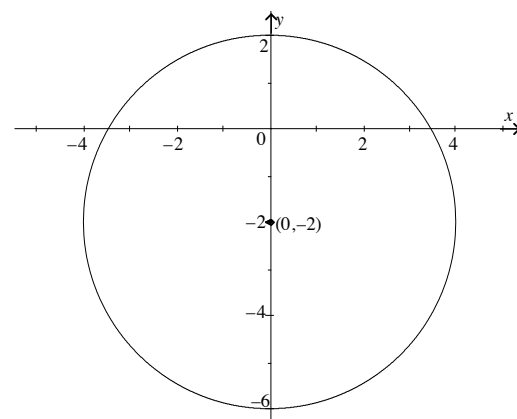
5.

a.



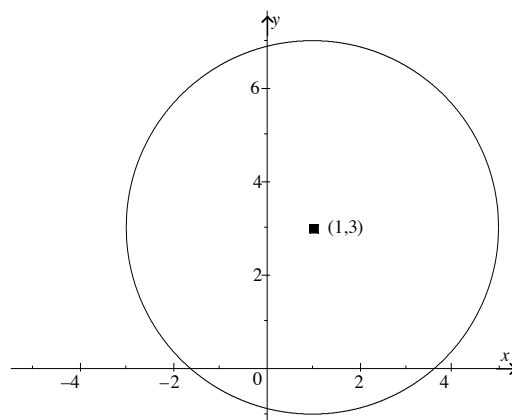
The graph of $x^2 + y^2 = 16$.

b.



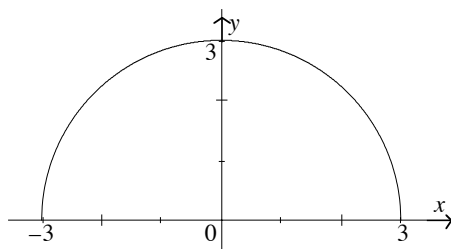
The graph of $x^2 + (y + 2)^2 = 16$.

c.



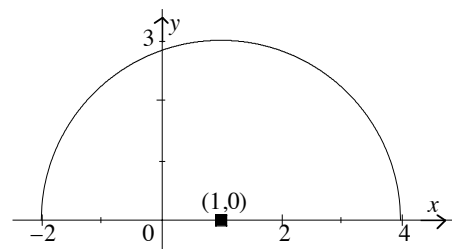
The graph of $(x - 1)^2 + (y - 3)^2 = 16$.

6. a.



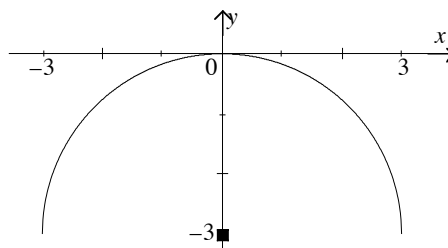
The graph of $y = \sqrt{9 - x^2}$.

b.



The graph of $y = \sqrt{9 - (x - 1)^2}$.

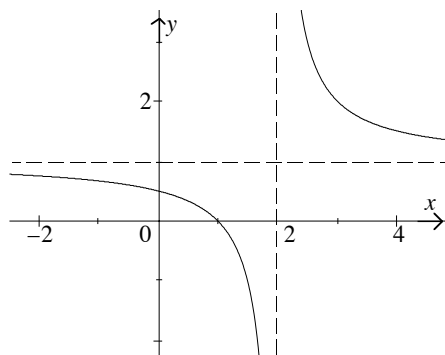
c.



The graph of $y = \sqrt{9 - x^2} - 3$.

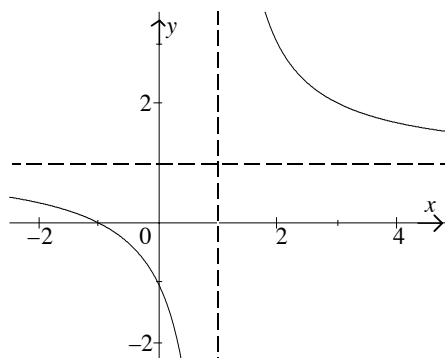
7.

$$\frac{1}{x-2} + 1 = \frac{1 + (x-2)}{x-2} = \frac{x-1}{x-2}$$



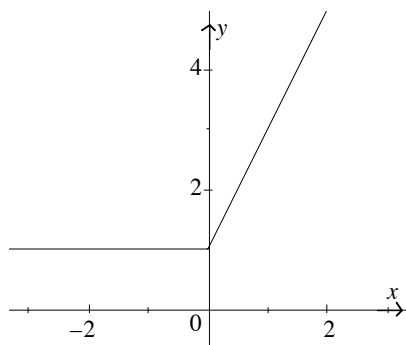
The graph of $y = \frac{x-1}{x-2}$.

8.



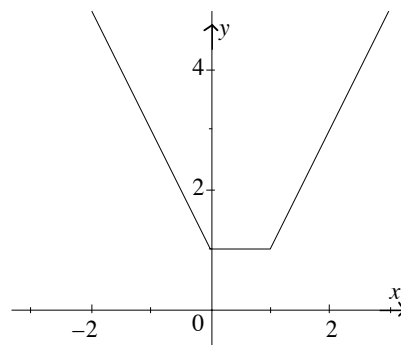
The graph of $y = \frac{x+1}{x-1}$.

9. a.



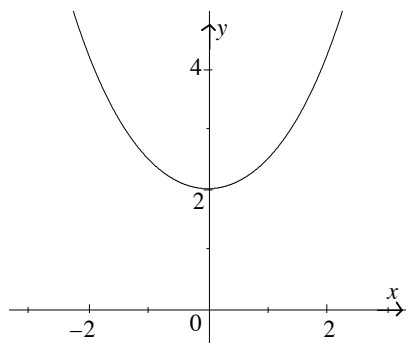
The graph of $y = |x| + x + 1$
for $-2 \leq x \leq 2$.

b.



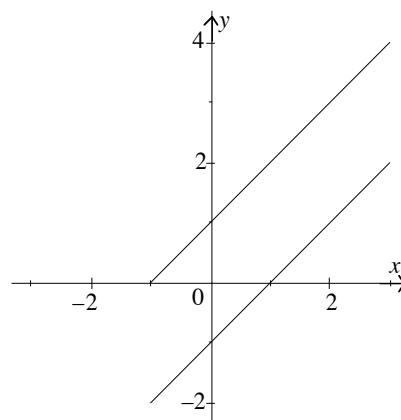
The graph of $y = |x| + |x - 1|$
for $-2 \leq x \leq 3$.

c.



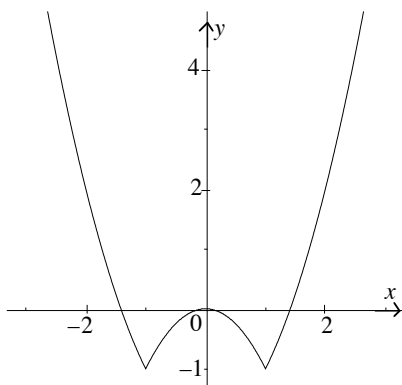
The graph of $y = 2^x + 2^{-x}$
for $-2 \leq x \leq 2$.

d.



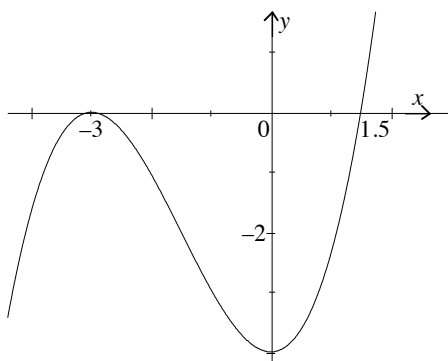
The graph of $|x - y| = 1$ for
 $-1 \leq x \leq 3$.

10.



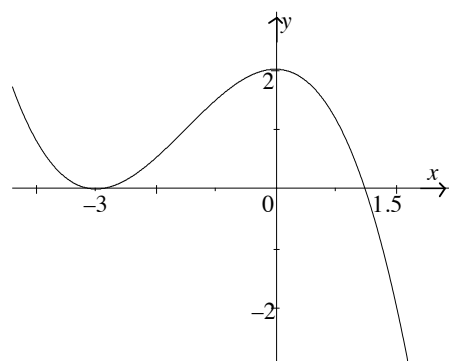
The graph of $f(x) = |x^2 - 1| - 1$.

11. a.



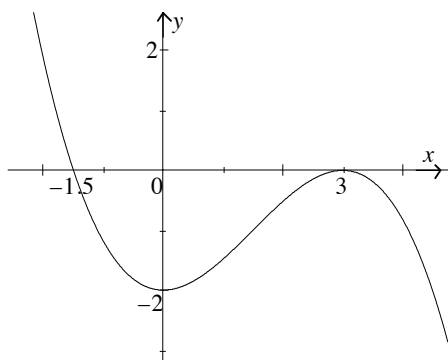
The graph of $y = 2f(x)$.

b.



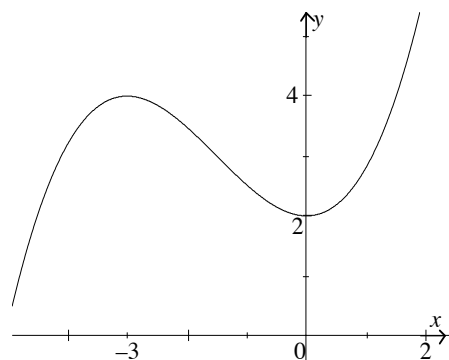
The graph of $y = -f(x)$.

c.



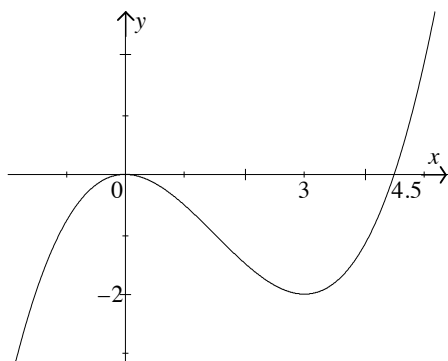
The graph of $y = f(-x)$.

d.



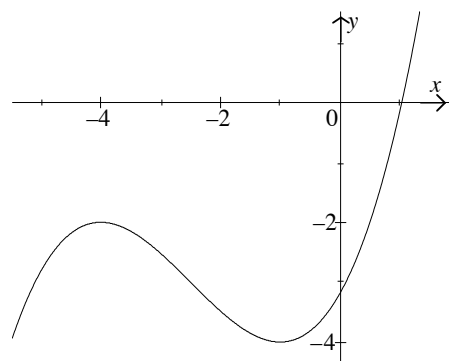
The graph of $y = f(x) + 4$.

e.



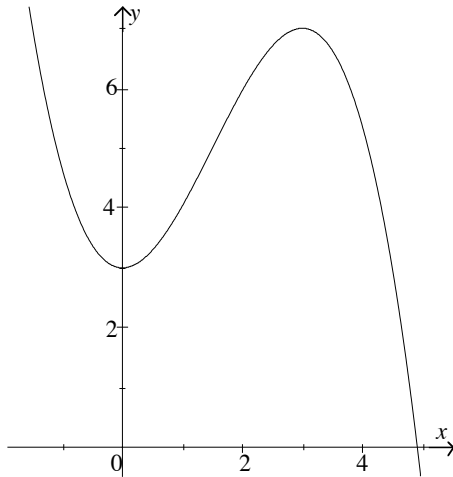
The graph of $y = f(x - 3)$.

f.



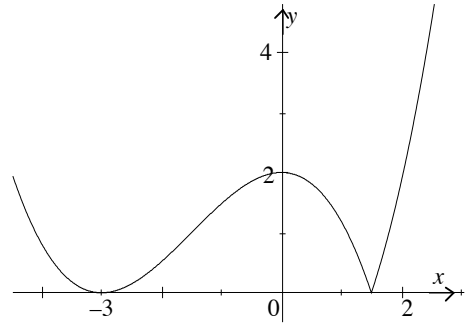
The graph of $y = f(x + 1) - 2$.

g.



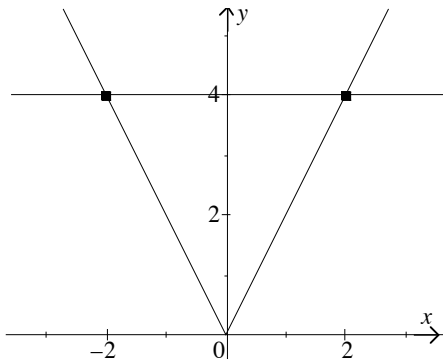
The graph of $y = 3 - 2f(x - 3)$.

h.



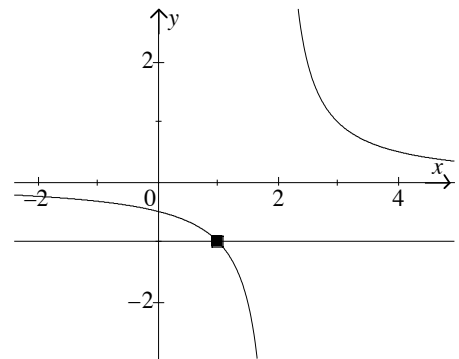
The graph of $y = |f(x)|$.

12. a.



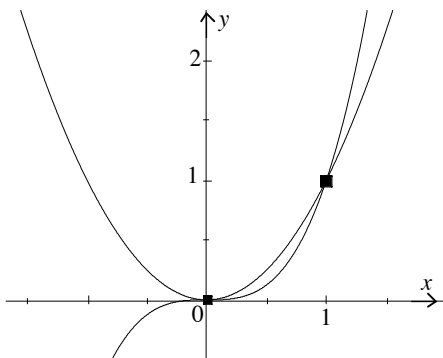
$x = -2$ and $x = 2$ are solutions of the equation $|2x| = 4$.

b.



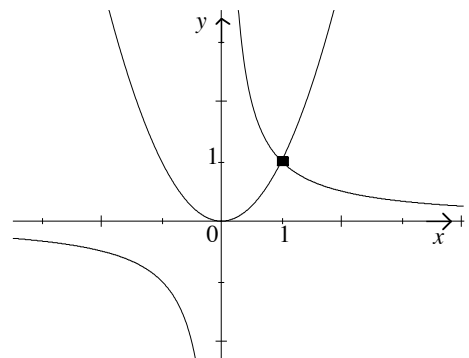
$x = 1$ is a solution of $\frac{1}{x-2} = -1$.

c.



$x = 0$ and $x = 1$ are solutions of the equation $x^3 = x^2$.

d.



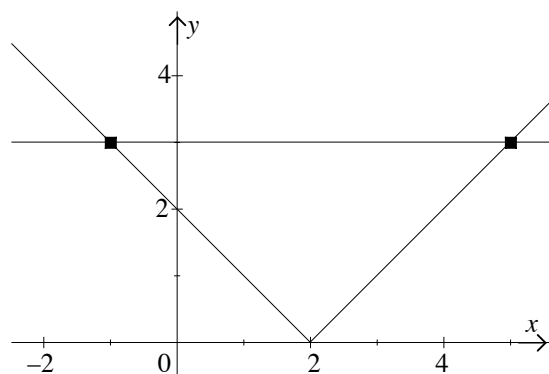
$x = 1$ is a solution of $x^2 = \frac{1}{x}$.

13.

a. For $x \geq 2$, $|x - 2| = x - 2 = 3$. Therefore $x = 5$ is a solution of the inequality. (Note that $x = 5$ is indeed ≥ 2 .)

For $x < 2$, $|x - 2| = -(x - 2) = -x + 2 = 3$. Therefore $x = -1$ is a solution. (Note that $x = -1$ is < 2 .)

b.

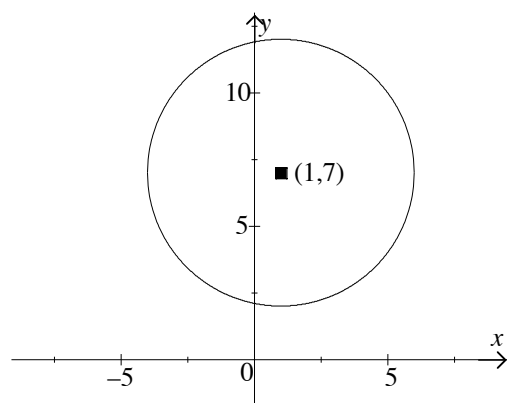


The points of intersection are $(-1, 3)$ and $(5, 3)$.

Therefore the solutions of $|x - 2| = 3$ are $x = -1$ and $x = 5$.

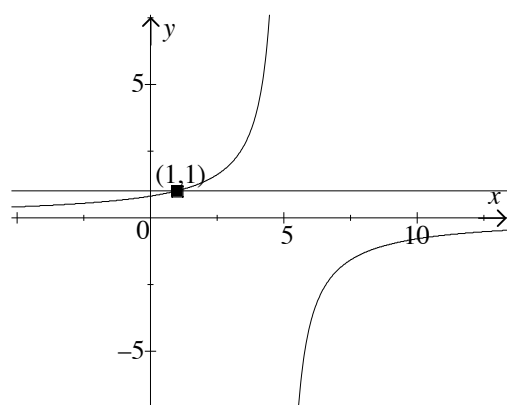
14. The parabolas intersect at $(2, 1)$.

15.



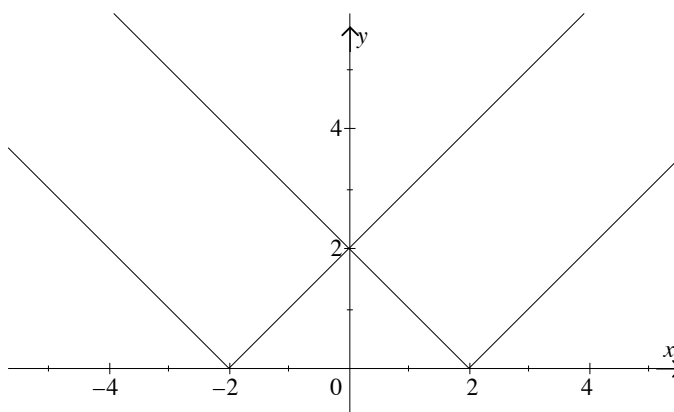
$y = k$ intersects the circle at two distinct points when $2 < k < 12$.

16.



The point of intersection is $(1, 1)$. Therefore the solution of $\frac{4}{5-x} = 1$ is $x = 1$.

17.



The point of intersection is $(0, 2)$. Therefore the solution of $|x - 2| = |x + 2|$ is $x = 0$.

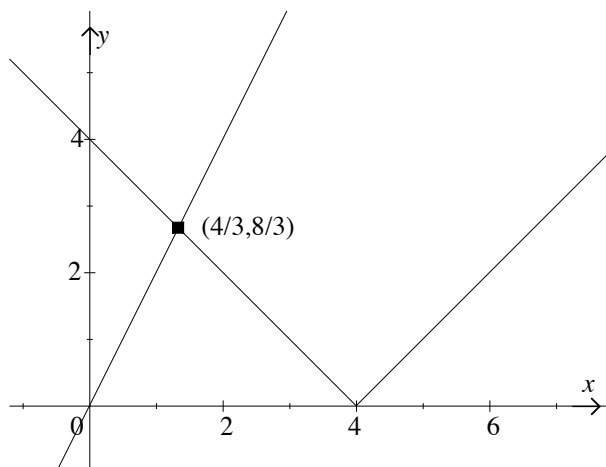
18. $n = -1$ or $n = 2$.

19. a. For $x \geq 4$, $|x - 4| = x - 4 = 2x$ when $x = -4$, but this does not satisfy the condition of $x \geq 4$ so is not a solution.

For $x < 4$, $|x - 4| = -x + 4 = 2x$ when $x = \frac{4}{3}$. $x = \frac{4}{3}$ is < 4 so is a solution.

Therefore, $x = \frac{4}{3}$ is a solution of $|x - 4| = 2x$.

b.



The graph of $y = |x - 4|$ and $y = 2x$ intersect at the point $(\frac{4}{3}, \frac{8}{3})$. So the solution of $|x - 4| = 2x$ is $x = \frac{4}{3}$.

2.11 Solutions

1. a. The domain is all real x , and the range is all real $y \geq -2$.

b. i $-2 < x < 0$ or $x > 2$

ii $x < -2$ or $0 < x < 2$

c. i $k < -2$

ii There is no value of k for which $f(x) = k$ has exactly one solution.

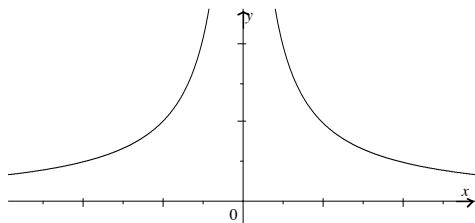
iii $k = 2$ or $k > 0$

iv $k = 0$

v $-2 < k < 0$

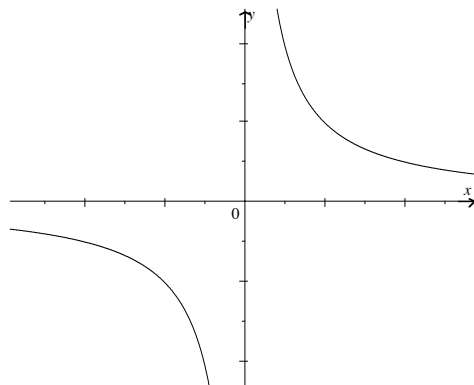
d. $y = f(x)$ is even

2. a.



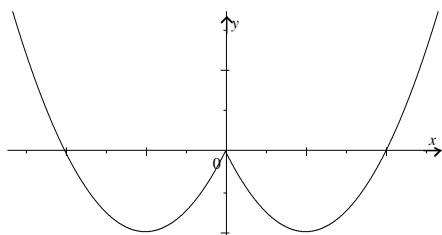
$y = f(x)$ is even.

b.



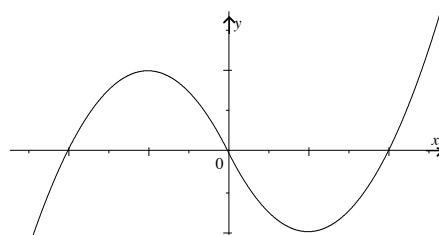
$y = f(x)$ is odd.

a.



$y = g(x)$ is even.

b.



$y = g(x)$ is odd.

3. a. even b. even c. neither d. odd e. odd
f. even g. even h. neither i. even j. even

4. a.

$$\begin{aligned} h(-x) &= f(-x) \cdot g(-x) \\ &= f(x) \cdot -g(x) \\ &= -f(x) \cdot g(x) \\ &= -h(x) \end{aligned}$$

Therefore h is odd.

b.

$$\begin{aligned} h(-x) &= (g(-x))^2 \\ &= (-(g(x)))^2 \\ &= (g(x))^2 \\ &= h(x) \end{aligned}$$

Therefore h is even.

c.

$$\begin{aligned}
 h(-x) &= \frac{f(-x)}{g(-x)} \\
 &= \frac{f(x)}{-g(x)} \\
 &= -\frac{f(x)}{g(x)} \\
 &= -h(x)
 \end{aligned}$$

Therefore h is odd.

d.

$$\begin{aligned}
 h(-x) &= f(-x) \cdot (g(-x))^2 \\
 &= f(x) \cdot (-g(x))^2 \\
 &= f(x) \cdot (g(x))^2 \\
 &= h(x)
 \end{aligned}$$

Therefore h is even.

5. If f is defined at $x = 0$

$$\begin{aligned}
 f(0) &= f(-0) && \text{(since } 0 = -0\text{)} \\
 &= -f(0) && \text{(since } f \text{ is odd)} \\
 2f(0) &= 0 && \text{(adding } f(0) \text{ to both sides)}
 \end{aligned}$$

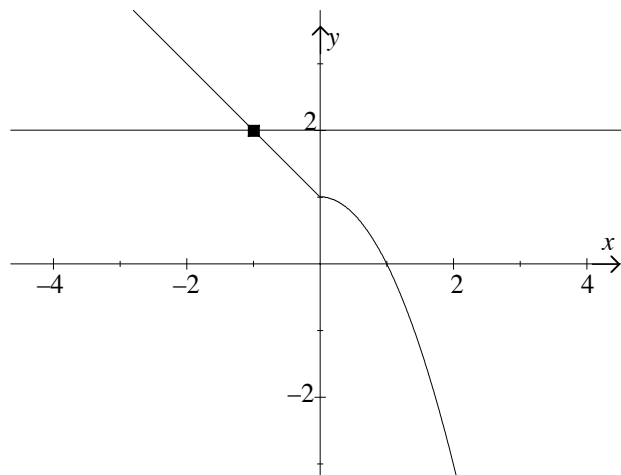
$$\text{Therefore } f(0) = 0.$$

3.2 Solutions

1. a. $2f(-1) + f(2) = 2(1 - (-1)) + (1 - (2)^2) = 4 + (-3) = 1.$

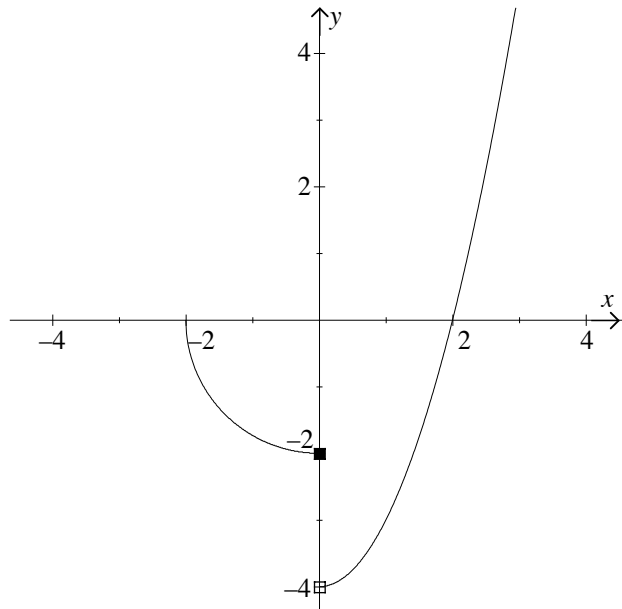
b. $f(a^2) = 1 - (a^2)^2 = 1 - a^4$ since $a^2 \geq 0$.

2. You can see from the graph below that there is one solution to $f(x) = 2$, and that this solution is at $x = -1$.



3.
$$g(x) = \begin{cases} \frac{1}{x+1} & \text{for } x < 1 \\ \sqrt{1-x^2} & \text{for } -1 \leq x \leq 1 \\ -1 & \text{for } x > 1 \end{cases}$$

4. a. The domain of f is all real $x \geq -2$.



b. The range of f is all real $y > -4$.

c. i $f(x) = 0$ when $x = -2$ or $x = 2$.

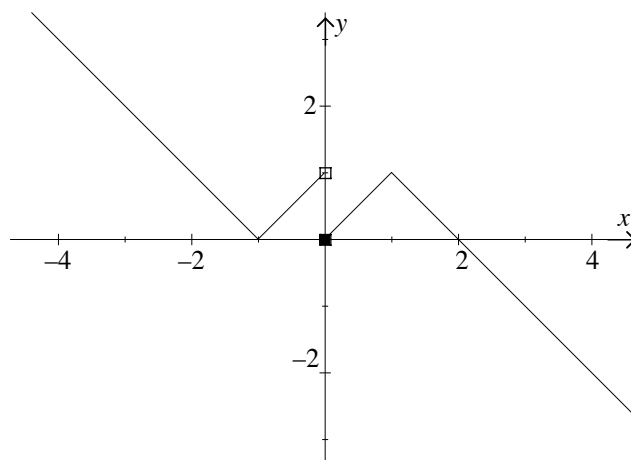
ii $f(x) = -3$ when $x = 1$.

d. i $f(x) = k$ has no solutions when $k \leq -4$.

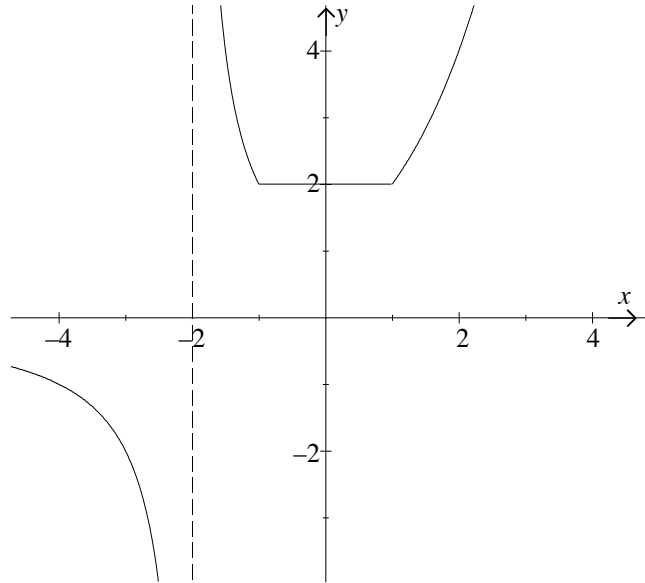
ii $f(x) = k$ has 1 solution when $-4 < k < -2$ or $k > 0$.

iii $f(x) = k$ has 2 solutions when $-2 \leq k \leq 0$.

5. Note that $f(0) = 0$.



6. The domain of g is all real x , $x \neq -2$.



The range of g is all real $y < 0$ or $y \geq 2$.

7. Note that there may be more than one correct solution.

a. Defining f as

$$f(x) = \begin{cases} x + 6 & \text{for } x \leq -3 \\ -x & \text{for } -3 < x < 0 \\ x & \text{for } 0 \leq x \leq 3 \\ -x + 6 & \text{for } x > 3 \end{cases}$$

gives a function describing the McMaths burgers' logo using 4 pieces.

b. Defining f as

$$f(x) = \begin{cases} x + 6 & \text{for } x \leq -3 \\ |x| & \text{for } -3 < x < 3 \\ -x + 6 & \text{for } x \geq 3 \end{cases}$$

gives a function describing the McMaths burgers' logo using 3 pieces.

c. Defining f as

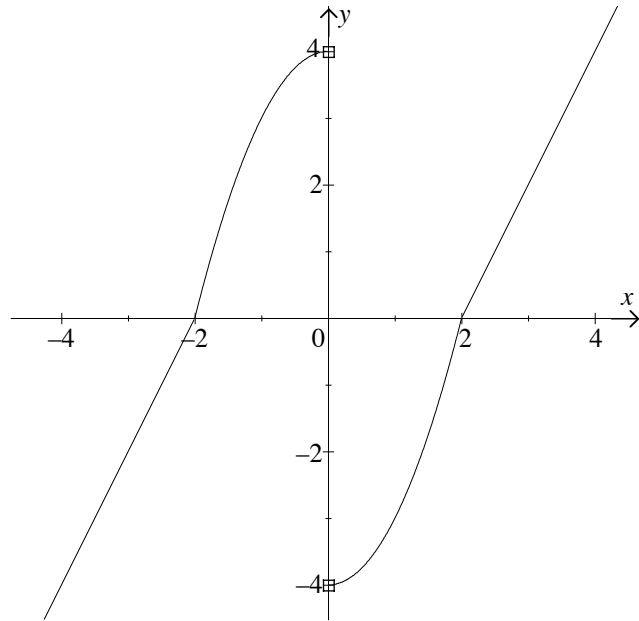
$$f(x) = \begin{cases} 3 - |x + 3| & \text{for } x \leq 0 \\ 3 - |x - 3| & \text{for } x > 0 \end{cases}$$

gives a function describing the McMaths burgers' logo using 2 pieces.

8. a. Here $a = 1$, $b = -4$, $c = 2$ and $d = -4$. So,

$$f(x) = \begin{cases} x^2 - 4 & \text{for } 0 < x \leq 2 \\ 2x - 4 & \text{for } x > 2 \end{cases}$$

b. Defining f to be an odd function for all real x , $x \neq 0$, we get

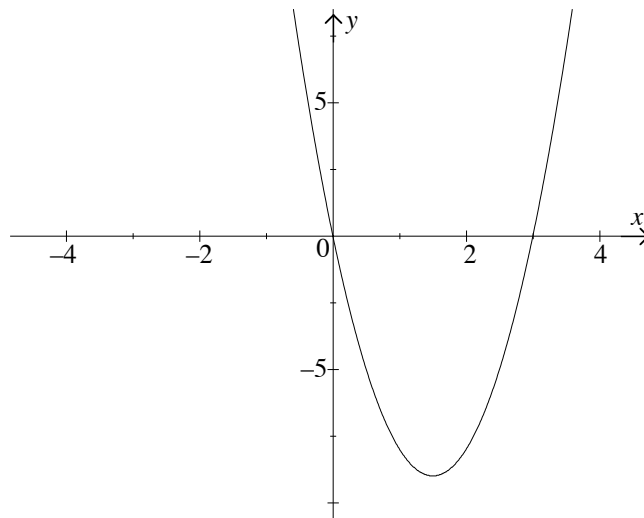


c. We can define f as follows

$$f(x) = \begin{cases} 2x + 4 & \text{for } x < -2 \\ 4 - x^2 & \text{for } -2 \leq x < 0 \\ x^2 - 4 & \text{for } 0 < x \leq 2 \\ 2x - 4 & \text{for } x > 2 \end{cases}$$

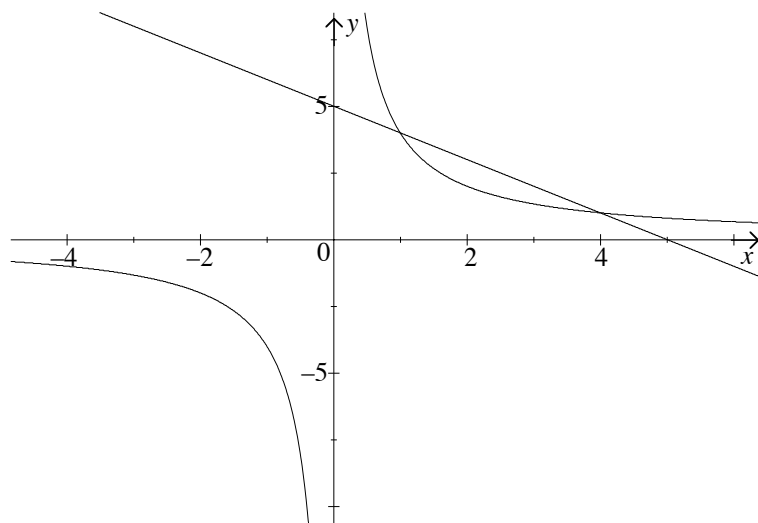
3.4 Solutions

1. a. $0 \leq x \leq 4$
b. $-3 < p \leq 1$
c. $x < -4$ or $-3 < x < 3$ or $x > 4$
2. a. The graph of $y = 4x(x - 3)$ is given below

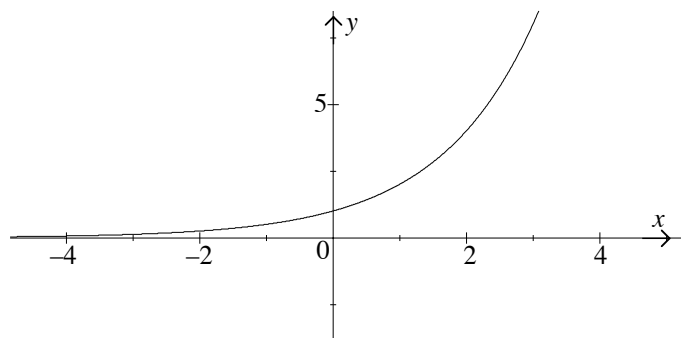


b. From the graph we see that $4x(x - 3) \leq 0$ when $0 \leq x \leq 3$.

3. a. The graphs $y = 5 - x$ and $y = \frac{4}{x}$ intersect at the points $(1, 4)$ and $(4, 1)$.
 b. The graphs of $y = 5 - x$ and $y = \frac{4}{x}$

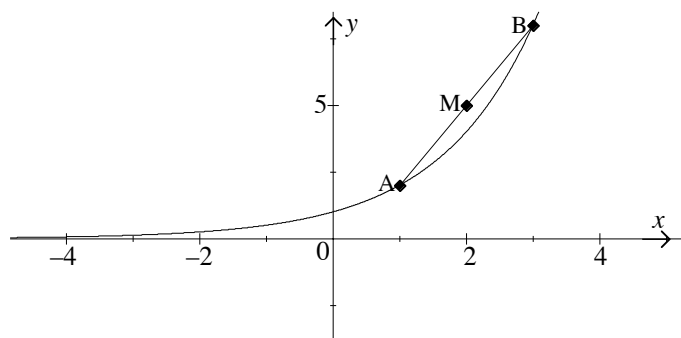


- c. The inequality is satisfied for $x < 0$ or $1 < x < 4$.
 4. a. The graph of $y = 2^x$.



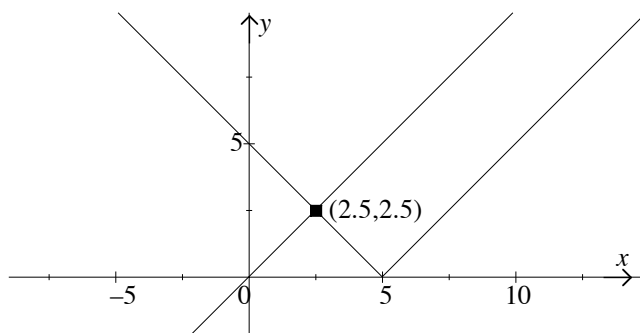
- b. $2^x < \frac{1}{2}$ when $x < -1$.
 c. The midpoint M of the segment AB has coordinates $(\frac{a+b}{2}, \frac{2^a+2^b}{2})$.
 Since the function $y = 2^x$ is concave up, the y -coordinate of M is greater than $f(\frac{a+b}{2})$. So,

$$\frac{2^a + 2^b}{2} > 2^{\frac{a+b}{2}}$$



5.

a.



b. $|x - 5| > x$ for all $x < 2.5$.

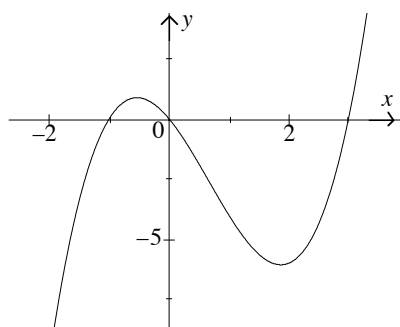
c. i $mx = |x - 5|$ has exactly two solutions when $0 < m < 1$.

ii $mx = |x - 5|$ has no solutions when $-1 < m < 0$.

6. $-1 \leq x \leq 3$

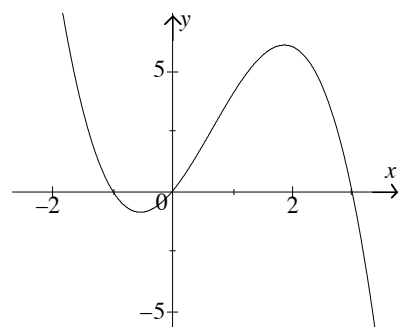
4.3 Solutions

1. a.



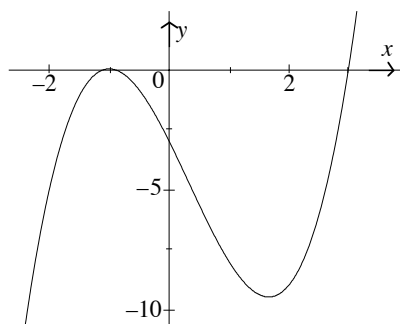
The graph of $P(x) = x(x+1)(x-3)$.

b.



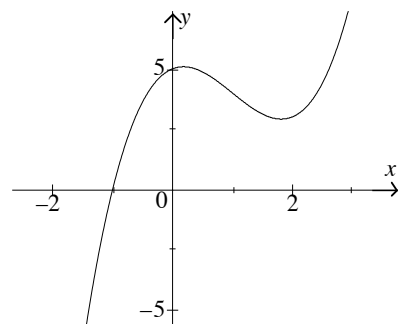
The graph of $P(x) = x(x+1)(3-x)$.

c.



The graph of $P(x) = (x+1)^2(x-3)$.

d.



The graph of $P(x) = (x+1)(x^2 - 4x + 5)$.

2. a. iv.

b. v.

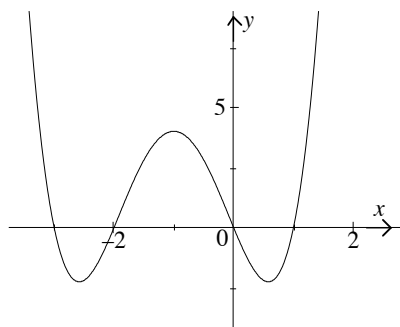
c. i.

d. iii.

e. ii.

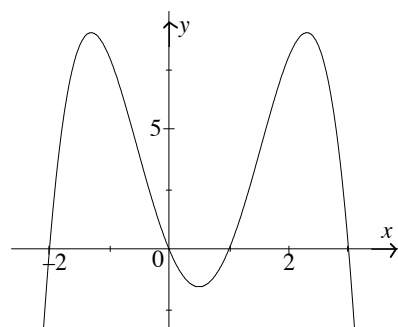
f. vi.

3. a.



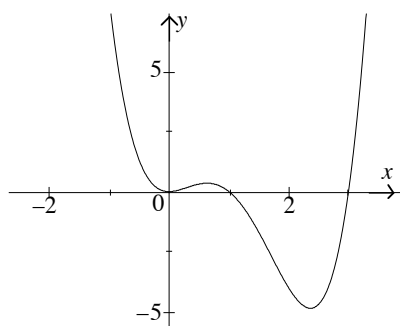
The graph of
 $P(x) = x(x-1)(x+2)(x+3)$.

b.



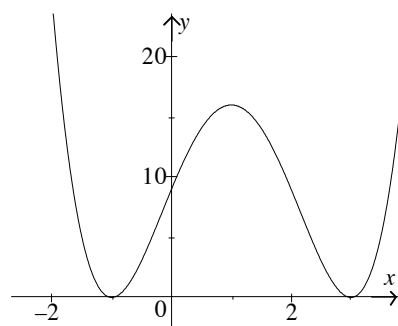
The graph of
 $P(x) = x(x-1)(x+2)(3-x)$.

c.



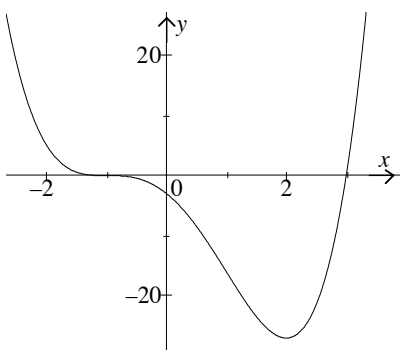
The graph of $P(x) = x^2(x-1)(x-3)$.

d.



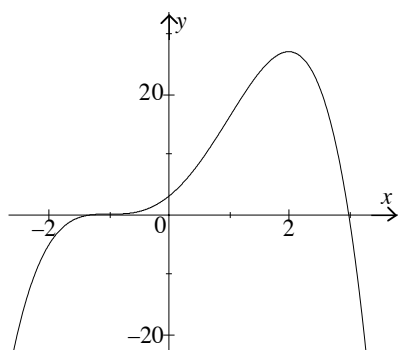
The graph of $P(x) = (x+1)^2(x-3)^2$.

e.



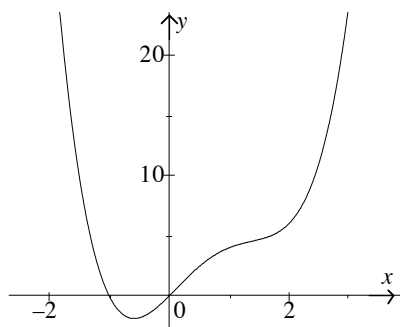
The graph of $P(x) = (x+1)^3(x-3)$.

f.



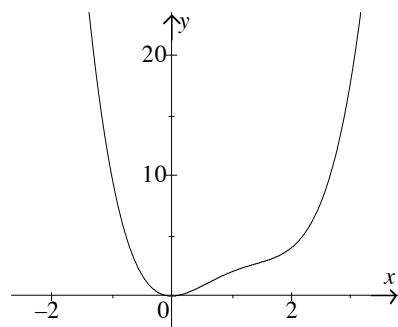
The graph of $P(x) = (x+1)^3(3-x)$.

g.



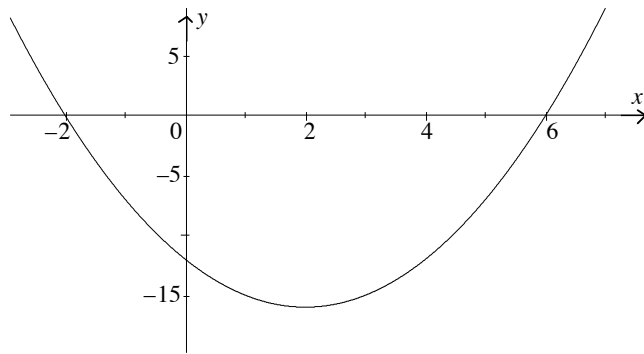
The graph of
 $P(x) = x(x+1)(x^2-4x+5)$.

h.



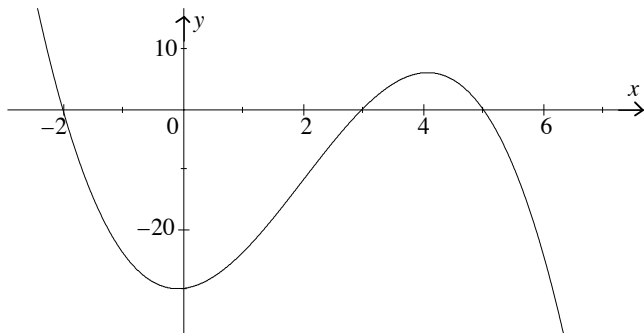
The graph of $P(x) = x^2(x^2-4x+5)$.

4. a.



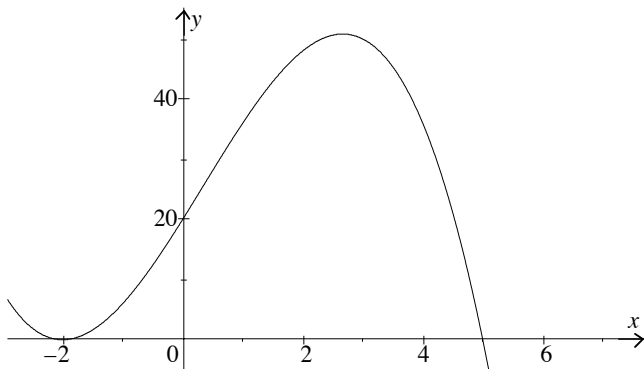
$$x^2 - 4x - 12 < 0 \text{ when } -2 < x < 6.$$

b.



$$(x + 2)(x - 3)(5 - x) > 0 \text{ when } x < -2 \text{ or } 3 < x < 5.$$

c.



$$(x + 2)^2(5 - x) > 0 \text{ when } x < 5.$$

5. $x^2 - 4x - 12 + k \geq 0$ for all real x when $k = 16$.

6. a. $P(x) = x(x - 4)$

b. $P(x) = -x(x - 4)$

c. $P(x) = x^2(x - 4)$

d. $P(x) = \frac{x^3(x-4)}{3}$

e. $P(x) = -x(x - 4)^2$

f. $P(x) = \frac{(x+4)(x-4)^2}{8}$

7. a. The roots of $f(x) = 0$ are $x = -2$, $x = 0$ and $x = 2$.

- b. $x = 2$ is the repeated root.
- c. The equation $f(x) = k$ has exactly 3 solutions when $k = 0$ or $k = 3.23$.
- d. $f(x) < 0$ when $-2 < x < 0$.
- e. The least possible degree of the polynomial $f(x)$ is 4.
- f. Since $f(0) = 0$, the constant in the polynomial is 0.
- g. $f(x) + k \geq 0$ for all real x when $k \geq 9.91$.

4.5 Solutions

1. a. Since $A(x) = (x - a)(x - b)$ is a polynomial of degree 2, the remainder $R(x)$ must be a polynomial of degree < 2 . So, $R(x)$ is a polynomial of degree ≤ 1 . That is, $R(x) = mx + c$ where m and c are constants. Note that if $m = 0$ the remainder is a constant.

- b. Let $P(x) = (x^2 - 5x + 6)Q(x) + (mx + c) = (x - 2)(x - 3)Q(x) + (mx + c)$.

Then

$$\begin{aligned} P(2) &= (0)(-1)Q(2) + (2m + c) \\ &= 2m + c \\ &= 4 \end{aligned}$$

and

$$\begin{aligned} P(3) &= (1)(0)Q(3) + (3m + c) \\ &= 3m + c \\ &= 9 \end{aligned}$$

Solving simultaneously we get that $m = 5$ and $c = -6$. So, the remainder is $R(x) = 5x - 6$.

- c. Let $P(x) = (x - a)(x - b)Q(x) + (mx + c)$.

Then

$$\begin{aligned} P(a) &= (0)(a - b)Q(a) + (ma + c) \\ &= am + c \\ &= a^2 \end{aligned}$$

and

$$\begin{aligned} P(b) &= (b - a)(0)Q(b) + (mb + c) \\ &= bm + c \\ &= b^2 \end{aligned}$$

Solving simultaneously we get that $m = a + b$ and $c = -ab$ provided $a \neq b$.

So, $R(x) = (a + b)x - ab$.

2. a.

$$2x^4 + 13x^3 + 18x^2 + x - 4 = (x^2 + 5x + 2)(2x^2 + 3x - 1) - 2$$

- b. Let α be a common zero of $f(x)$ and $g(x)$. That is, $f(\alpha) = 0$ and $g(\alpha) = 0$.

Then since $f(x) = g(x)q(x) + r(x)$ we have

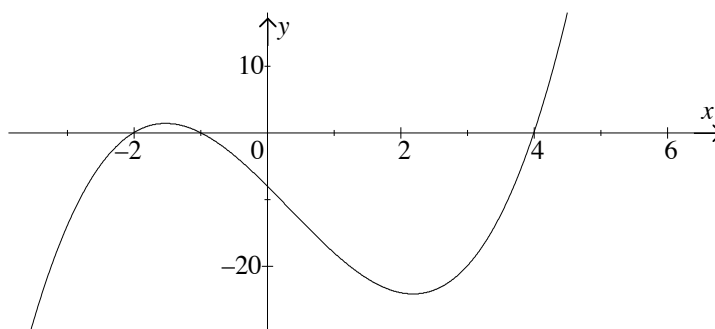
$$\begin{aligned} f(\alpha) &= g(\alpha)q(\alpha) + r(\alpha) \\ &= (0)q(\alpha) + r(\alpha) \quad \text{since } g(\alpha) = 0 \\ &= r(\alpha) \\ &= 0 \quad \text{since } f(\alpha) = 0 \end{aligned}$$

But, from part **b**. $r(x) = -2$ for all values of x , so we have a contradiction.

Therefore, $f(x)$ and $g(x)$ do not have a common zero.

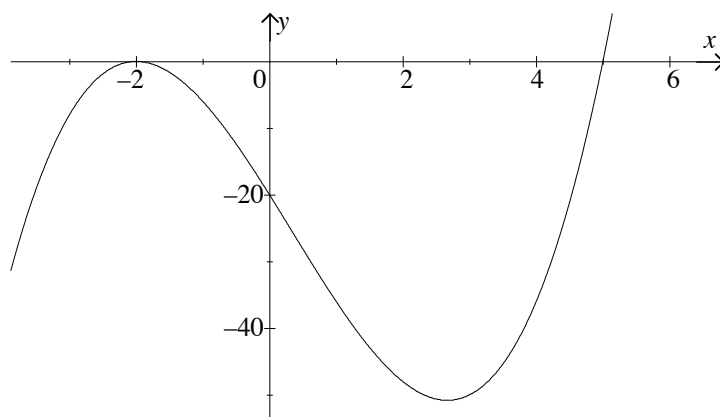
This is an example of a proof by contradiction.

3. a. i $P(x) = x^3 - x^2 - 10x - 8 = (x+1)(x+2)(x-4)$
 ii $x = -1$, $x = -2$ and $x = 4$ are solutions of $P(x) = 0$.
 iii



The graph of $P(x) = x^3 - x^2 - 10x - 8$.

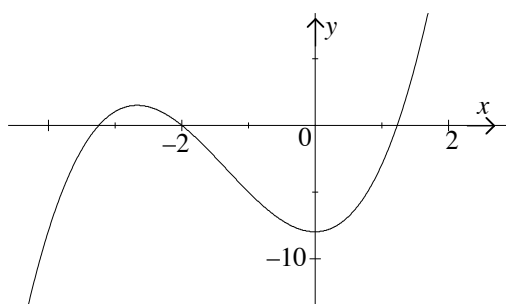
- b. i $P(x) = x^3 - x^2 - 16x - 20 = (x+2)^2(x-5)$.
 ii $x = -2$ and $x = 5$ are solutions of $P(x) = 0$. $x = -2$ is a double root.
 iii



The graph of $P(x) = x^3 - x^2 - 16x - 20$.

- c. i $P(x) = x^3 + 4x^2 - 8 = (x+2)(x^2 + 2x - 4) = (x+2)(x - (-1 + \sqrt{5}))(x - (-1 - \sqrt{5}))$
 ii $x = -2$, $x = -1 + \sqrt{5}$ and $x = -1 - \sqrt{5}$ are solutions of $P(x) = 0$.

iii



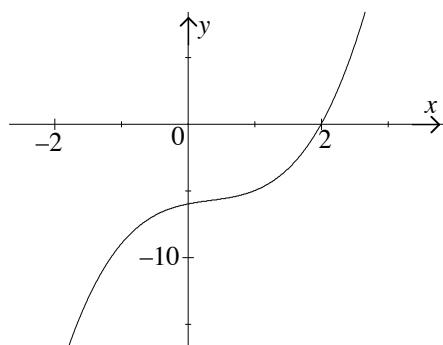
The graph of $P(x) = x^3 + 4x^2 - 8$.

The zeros are $x = -2$, $x = -1 + \sqrt{5}$ and $x = -1 - \sqrt{5}$.

d. i $P(x) = x^3 - x^2 + x - 6 = (x - 2)(x^2 + x + 3)$. $x^2 + x + 3 = 0$ has no real solutions.

ii $x = 2$ is the only real solution of $P(x) = 0$.

iii



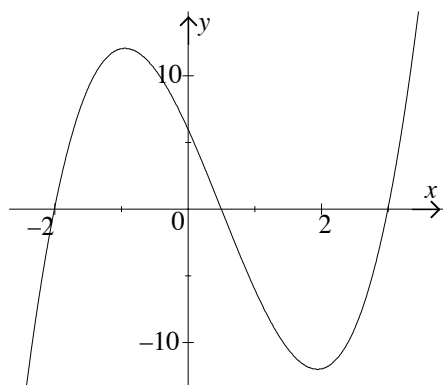
The graph of $P(x) = x^3 - x^2 + x - 6$.

There is only one real zero at $x = 2$.

e. i $P(x) = 2x^3 - 3x^2 - 11 + 6 = (x + 2)(x - 3)(2x - 1)$.

ii $x = -2$, $x = \frac{1}{2}$ and $x = 3$ are solutions of $P(x) = 0$.

iii



The graph of $P(x) = 2x^3 - 3x^2 - 11 + 6$.

Chapter four:

Lecture notes on relations and functions

1. RELATIONS

1.1. The idea of a relation. Let X and Y be two sets. We would like to formalize the idea of a **relation** between X and Y . Intuitively speaking, this is a well-defined “property” R such that given any $x \in X$ and $y \in Y$, either x bears the property R to y , or it doesn’t (and not both!). Some important examples:

Example 1.1. Let X be a set of objects and let Y be a set of sets. Then “membership” is a relation R from X to Y : i.e., we have xRy if $x \in y$.

Example 1.2. Let S be a set, and let $X = Y = 2^S$, the power set of S (recall that this is the set of all subsets of S). Then containment, $A \subseteq B$ is a relation between X and Y . (Proper containment, $A \subset B$, is also a relation.)

Example 1.3. Let $X = Y$. Then equality is a relation from X to Y : we say xRy iff $x = y$. Also inequality is a relation between X and Y : we say xRy iff $x \neq y$.

Example 1.4. Let $X = Y = \mathbb{R}$. Then $\leq, <, \geq, >$ are relations between \mathbb{R} and \mathbb{R} .

Example 1.5. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function. Then we can define a relation from \mathbb{R} to \mathbb{R} , by xRy if and only if $y = f(x)$.

Date: April 15, 2016.

Example 1.6. Let $X = Y = \mathbb{Z}$. Then divisibility is a relation between \mathbb{Z} and \mathbb{Z} : we say xRy if $x \mid y$.

Example 1.7. Let $X = Y = \mathbb{Z}$. Then “having the same parity” is a relation between \mathbb{Z} and \mathbb{Z} .

In many of the above examples we have $X = Y$. This will often (but certainly not always!) be the case, and when it is we may speak of relations **on** \mathbf{X} .

1.2. The formal definition of a relation.

We still have not given a formal definition of a relation between sets X and Y . In fact the above way of thinking about relations is easily formalized, as was suggested in class by Adam Osborne: namely, we can think of a relation R as a function from $X \times Y$ to the two-element set $\{\text{TRUE}, \text{FALSE}\}$. In other words, for $(x, y) \in X \times Y$, we say that xRy if and only if $f((x, y)) = \text{TRUE}$.

This is a great way of thinking about relations. It has however one foundational drawback: it makes the definition of a relation depend on that of a function, whereas the standard practice for about one hundred years is the reverse: we want to define a function as a special kind of relation (c.f. Example 5 above). The familiar correspondence between logic and set theory leads us to the official definition:

Definition: A relation R between two sets X and Y is simply a **subset** of the Cartesian product $X \times Y$, i.e., a collection of ordered pairs (x, y) .

(Thus we have replaced the basic logical dichotomy “TRUE/FALSE” with the basic set-theoretic dichotomy “is a member of/ is not a member of”.) Note that this new definition has some geometric appeal: we are essentially identifying a relation R with its *graph* in the sense of precalculus mathematics.

We take advantage of the definition to adjust the terminology: rather than speaking (slightly awkwardly) of relations “from X to Y ” we will now speak of relations **on** $\mathbf{X} \times \mathbf{Y}$. When $X = Y$ we may (but need not!) speak of relations **on** \mathbf{X} .

Example 1.8. Any curve in \mathbb{R}^2 defines a relation on $\mathbb{R} \times \mathbb{R}$. E.g. the unit circle

$$x^2 + y^2 = 1$$

is a relation in the plane: it is just a set of ordered pairs.

1.3. Basic terminology and further examples.

Let X, Y be sets. We consider the set of all relations on $X \times Y$ and denote it by $\mathcal{R}(X, Y)$. According to our formal definition we have

$$\mathcal{R}(X, Y) = 2^{X \times Y},$$

i.e., the set of all subsets of the Cartesian product $X \times Y$.

Example 1.9. a) Suppose $X = \emptyset$. Then $X \times Y = \emptyset$ and $\mathcal{R}(X \times Y) = 2^\emptyset = \{\emptyset\}$. That is: if X is empty, then the set of ordered pairs (x, y) for $x \in X$ and $y \in Y$ is empty, so there is only one relation: the empty relation.

b) Suppose $Y = \emptyset$. Again $X \times Y = \emptyset$ and the discussion is the same as above.

Example 1.10. a) Suppose $X = \{\bullet\}$ consists of a single element. Then $X \times Y = \{(\bullet, y) \mid y \in Y\}$; in other words, $X \times Y$ is essentially just Y itself, since the first coordinate is always the same. Thus a relation R on $X \times Y$ corresponds to a subset of Y : formally, the set of all $y \in Y$ such that $\bullet Ry$.

b) Suppose $Y = \{\bullet\}$ consists of a single element. The discussion is analogous to that of part a), and relations on $X \times Y$ correspond to subsets of X .

Example 1.11. Suppose X and Y are finite sets, with $\#X = m$ and $\#Y = n$. Then $\mathcal{R}(X, Y) = 2^{X \times Y}$ is finite, of cardinality

$$\#2^{X \times Y} = 2^{\#X \times Y} = 2^{\#X \cdot \#Y} = 2^{mn}.$$

The function 2^{mn} grows rapidly with both m and n , and the upshot is that if X and Y are even moderately large finite sets, the set of all relations on $X \times Y$ is very large. For instance if $X = \{a, b\}$ and $Y = \{1, 2\}$ then there are $2^{2 \cdot 2} = 16$ relations on $X \times Y$. It is probably a good exercise for you to write them all down. However, if $X = \{a, b, c\}$ and $Y = \{1, 2, 3\}$ then there are $2^{3 \cdot 3} = 512$ relations on $X \times Y$, and – with apologies to the Jackson 5? – it is less easy to write them all down.

Exercise 1.1. Let X and Y be nonempty sets, at least one of which is infinite. Show: $\mathcal{R}(X, Y)$ is infinite.

Given two relations R_1 and R_2 between X and Y , it makes sense to say that $R_1 \subseteq R_2$: this means that R_1 is “stricter” than R_2 or that R_2 is “more permissive” than R_1 . This is a very natural idea: for instance, if X is the set of people in the world, R_1 is the brotherhood relation – i.e., $(x, y) \in R_1$ iff x and y are brothers – and R_2 is the sibling relation – i.e., $(x, y) \in R_2$ iff x and y are siblings – then $R_1 \subsetneq R_2$: if x and y are brothers then they are also siblings, but not conversely.

Among all elements of $\mathcal{R}(X, Y)$, there is one relation R_\emptyset which is the strictest of all, namely $R_\emptyset = \emptyset$:¹ that is, for no $(x, y) \in X \times Y$ do we have $(x, y) \in R_\emptyset$. Indeed $R_\emptyset \subset R$ for any $R \in \mathcal{R}(X, Y)$. At the other extreme, there is a relation which is the most permissive, namely $R_{X \times Y} = X \times Y$ itself: that is, for all $(x, y) \in X \times Y$ we have $(x, y) \in R_{X \times Y}$. And indeed $R \subset R_{X \times Y}$ for any $R \in \mathcal{R}(X, Y)$.

Example 1.12. Let $X = Y$. The equality relation $R = \{(x, x) \mid x \in X\}$ can be thought of geometrically as the diagonal of $X \times Y$.

The **domain**² of a relation $R \subseteq X \times Y$ is the set of $x \in X$ such that there exists $y \in Y$ with $(x, y) \in R$. In other words, it is the set of all elements in X which relate to at least one element of Y .

Example 1.13. The circle relation $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ has domain $[-1, 1]$.

Given a relation $R \subseteq X \times Y$, we can define the **inverse relation** $R^{-1} \subseteq Y \times X$ by interchanging the order of the coordinates. Formally, we put

$$R^{-1} = \{(y, x) \in Y \times X \mid (x, y) \in R\}.$$

Geometrically, this corresponds to reflecting across the line $y = x$.

¹The notation here is just to emphasize that we are viewing \emptyset as a relation on $X \times Y$.

²I don't like this terminology. But it is used in the course text, and it would be confusing to change it.

Example 1.14. Consider the relation $R \subset \mathbb{R} \times \mathbb{R}$ attached to the function $f(x) = x^2$:

$$R = \{(x, x^2) \mid x \in \mathbb{R}\}.$$

The graph of this relation is an upward-opening parabola: it can also be described by the equation $y = x^2$. The inverse relation R^{-1} is $\{(x^2, x) \mid x \in \mathbb{R}\}$, which corresponds to the equation $x = y^2$ and geometrically is a parabola opening rightward. Note that the domain of the original relation R is \mathbb{R} , whereas the domain of R^{-1} is $[0, \infty)$. Moreover, R^{-1} is not a function, since some values of x relate to more than one y -value: e.g. $(1, 1)$ and $(1, -1)$ are both in R^{-1} .

Example 1.15. Consider the relation attached to the function $f(x) = x^3$: namely

$$R = \{(x, x^3) \mid x \in \mathbb{R}\}.$$

This relation is described by the equation $y = x^3$; certainly it is a function, and its domain is \mathbb{R} . Consider the inverse relation

$$R^{-1} = \{(x^3, x) \mid x \in \mathbb{R}\},$$

which is described by the equation $x = y^3$. Since every real number has a unique real cube root, this is equivalent to $y = x^{\frac{1}{3}}$. Thus this time R^{-1} is again a function, and its domain is \mathbb{R} .

Later we will study functions in detail and one of our main goals will be to understand the difference between Examples 1.14 and 1.15.

1.4. Properties of relations.

Let X be a set. We now consider various properties that a relation R on X – i.e., $R \subset X \times X$ may or may not possess.

Reflexivity: For all $x \in X$, $(x, x) \in R$.

In other words, each element of X bears relation R to itself. Another way to say this is that the relation R contains the equality relation on X .

Exercise 1.2. Which of the relations in Examples 1.1 through 1.15 are reflexive?

Anti-reflexivity: For all $x \in X$, $(x, x) \notin R$.

Certainly no relation on X is both reflexive and anti-reflexive (except in the silly case $X = \emptyset$ when both properties hold vacuously). However, notice that a relation need not be either reflexive or anti-reflexive: if there are $x, y \in X$ such that $(x, x) \in R$ and $(y, y) \notin R$, then neither property holds.

Symmetry: For all $x, y \in X$, if $(x, y) \in R$, then $(y, x) \in R$.

Again, this has a geometric interpretation in terms of symmetry across the diagonal $y = x$. For instance, the relation associated to the function $y = \frac{1}{x}$ is symmetric since interchanging x and y changes nothing, whereas the relation associated to the function $y = x^2$ is not. (Looking ahead a bit, a function $y = f(x)$ is symmetric iff it coincides with its own inverse function.)

Exercise 1.3. Which of the relations in Examples 1.1 through 1.15 are symmetric?

Example 1.16. Let V be a set. A **(simple, loopless, undirected) graph** – in the sense of graph theory, not graphs of functions! – is given by a relation E on V which is irreflexive and symmetric. Thus: for $x, y \in V$, we say that x and y are **adjacent** if $(x, y) \in E$. Moreover x is never adjacent to itself, and the adjacency of x and y is a property of the unordered pair $\{x, y\}$: if x is adjacent to y then y is adjacent to x .

Anti-Symmetry: for all $x, y \in X$, if $(x, y) \in R$ and $(y, x) \in R$, then $x = y$.

Exercise 1.4. Which of the relations in Examples 1.1 through 1.16 are anti-symmetric?

Transitivity: for all $x, y, z \in X$, if $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$.

“Being a parent of” is not transitive, but “being an ancestor of” is transitive.

Exercise 1.5. Which of the relations in Examples 1.1 through 1.15 are transitive?

Worked Exercise 1.6.

Let R be a relation on X . Show the following are equivalent:

- (i) R is both symmetric and anti-symmetric.
- (ii) R is a subrelation of the equality relation.

Solution: Suppose that we have a relation R on X which is both symmetric and anti-symmetric. Then, for all $x, y \in R$, if $(x, y) \in R$, then by symmetry we have also $(y, x) \in R$, and then by anti-symmetry we have $x = y$. Thus we’ve shown that if (i) holds, the only possible elements $(x, y) \in R$ are those of the form (x, x) , which means that R is a subrelation of the equality relation. Conversely, if R is a subrelation of equality and $(x, y) \in R$, then $y = x$, so $(y, x) \in R$. Similarly, if $(x, y) \in R$ and $(y, x) \in R$ then $x = y$. So R is both symmetric and anti-symmetric.

Now we make two further definitions of relations with possess certain combinations of these basic properties. The first is the most important definition in this section.

An **equivalence relation** on a set X is a relation on X which is reflexive, symmetric and transitive.

A **partial ordering** on a set X is a relation on X which is reflexive, anti-symmetric and transitive.

Exercise 1.7. Which of the relations in Examples 1.1 through 1.16 are equivalence relations? Which are partial orderings?

We often denote equivalence relations by a tilde – $x \sim y$ – and read $x \sim y$ as “ x is equivalent to y ”. For instance, the relation “having the same parity” on \mathbb{Z} is an equivalence relation, and $x \sim y$ means that x and y are both even or both odd. Thus it serves to group the elements of \mathbb{Z} into subsets which share some common property. In this case, all the even numbers are being grouped together and all the odd numbers are being grouped together. We will see shortly that this is a characteristic property of equivalence relations: every equivalence relation on a set X determines a **partition** on X and conversely, given any partition on X we can define an equivalence relation.

The concept of a partial ordering should be regarded as a “generalized less than

or equal to" relation. Perhaps the best example is the containment relation \subseteq on the power set $\mathcal{P}(S)$ of a set S . This is a very natural way of regarding one set as "bigger" or "smaller" than another set. Thus the insight here is that containment satisfies many of the formal properties of the more familiar \leq on numbers. However there is one property of \leq on numbers that does not generalize to \subseteq (and hence not to an arbitrary partial ordering): namely, given any two real numbers x, y we must have either $x \leq y$ or $y \leq x$. However for sets this does not need to be the case (unless S has at most one element). For instance, in the power set of the positive integers, we have $A = \{1\}$ and $B = \{2\}$, so neither is it true that $A \subseteq B$ or that $B \subseteq A$. This is a much stronger property of a relation:

Totality: For all $x, y \in X$, either $(x, y) \in R$ or $(y, x) \in R$.

A **total ordering** (or **linear ordering**) on a set X is a partial ordering satisfying dichotomy.

Example 1.17. *The relation \leq on \mathbb{R} is a total ordering.*

There is an entire branch of mathematics – **order theory** – devoted to the study of partial orderings.³ In my opinion order theory gets short shrift in the standard mathematics curriculum (especially at the advanced undergraduate and graduate levels): most students learn only a few isolated results which they apply frequently but with little context or insight. Unfortunately we are not in a position to combat this trend: partial and total orderings will get short shrift here as well!

1.5. Partitions and Equivalence Relations.

Let X be a set, and let \sim be an equivalence relation on X .

For $x \in X$, we define the **equivalence class of x** as

$$[x] = \{y \in X \mid y \sim x\}.$$

For example, if \sim is the relation "having the same parity" on \mathbb{Z} , then

$$[2] = \{\dots, -4, -2, 0, 2, 4, \dots\},$$

i.e., the set of all even integers. Similarly

$$[1] = \{\dots -3, -1, 1, 3, \dots\}$$

is the set of all odd integers. But an equivalence class in general has many "representatives". For instance, the equivalence class $[4]$ is the set of all integers having the same parity as 4, so is again the set of all even integers: $[4] = [2]$. More generally, for any even integer n , we have $[n] = [0]$ and for any odd integer n we have $[n] = [1]$. Thus in this case we have partitioned the integers into two subsets: the even integers and the odd integers.

We claim that given any equivalence relation \sim on a set X , the set $\{[x] \mid x \in X\}$ forms a partition of X . Before we proceed to demonstrate this, observe that we are now strongly using our convention that there is no "multiplicity" associated to membership in a set: e.g. the sets $\{4, 2 + 2, 1^1 + 3^0 + 2^1\}$ and $\{4\}$ are equal. The

³For instance, there is a journal called **Order**, in which a paper of mine appears.

above representation $\{[x] \mid x \in X\}$ is highly redundant: for instance in the above example we are writing down the set of even integers and the set of odd integers infinitely many times, but it only “counts once” in order to build the set of subsets which gives the partition.

With this disposed of, the verification that $\mathcal{P} = \{[x] \mid x \in X\}$ gives a partition of X comes down to recalling the definition of a partition and then following our noses. There are three properties to verify:

- (i) That every element of \mathcal{P} is nonempty. Indeed, the element $[x]$ is nonempty because it contains x ! This is by reflexivity: $x \sim x$, so $x \in \{y \in X \mid y \sim x\}$.
- (ii) That the union of all the elements of \mathcal{P} is all of X . But again, the union is indexed by the elements x of X , and we just saw that $x \in [x]$, so every x in X is indeed in at least one element of \mathcal{P} .
- (iii) Finally, we must show that if $[x] \cap [y] \neq \emptyset$, then $[x] = [y]$: i.e., any two elements of \mathcal{P} which have a common element must be the same element. So suppose that there exists $z \in [x] \cap [y]$. Writing this out, we have $z \sim x$ and $z \sim y$. By symmetry, we have $y \sim z$; from this and $z \sim x$, we deduce by transitivity that $y \sim x$, i.e., $y \in [x]$. We claim that it follows from this that $[y] \subset [x]$. To see this, take any $w \in [y]$, so that $w \sim y$. Since $w \sim x$, we conclude $w \sim x$, so $w \in [x]$. Rerunning the above argument with the roles of x and y interchanged we get also that $[y] \subset [x]$, so $[x] = [y]$. This completes the verification.

Note that the key fact underlying the proof was that any two equivalence classes $[x]$ and $[y]$ are either disjoint or coincident. Note also that we did indeed use all three properties of an equivalence relation.

Now we wish to go in the other direction. Suppose X is a set and $\mathcal{P} = \{U_i\}_{i \in I}$ is a partition of X (here I is just an index set). We can define an equivalence relation \sim on X as follows: we say that $x \sim y$ if there exists $i \in I$ such that $x, y \in U_i$. In other words, we are decreeing x and y to be equivalent exactly when they lie in the same “piece” of the partition. Let us verify that this is an equivalence relation. First, let $x \in X$. Then, since \mathcal{P} is a partition, there exists some $i \in I$ such that $x \in U_i$, and then x and x are both in U_i , so $x \sim x$. Next, suppose that $x \sim y$: this means that there exists $i \in I$ such that x and y are both in U_i ; but then sure enough y and x are both in U_i (“and” is commutative!), so $y \sim x$. Similarly, if we have x, y, z such that $x \sim y$ and $y \sim z$, then there exists i such that x and y are both in U_i and a possibly different index j such that y and z are both in U_j . But since $y \in U_i \cap U_j$, we must have $U_i = U_j$ so that x and z are both in $U_i = U_j$ and $x \sim z$.

Moreover, the processes of passing from an equivalence relation to a partition and from a partition to an equivalence relation are mutually inverse: if we start with an equivalence relation R , form the associated partition $\mathcal{P}(R)$, and then form the associated equivalence relation $\sim (\mathcal{P}(R))$, then we get the equivalence relation R that we started with, and similarly in the other direction.

1.6. Examples of equivalence relations.

Example 1.18. (Congruence modulo n) Let $n \in \mathbb{Z}^+$. There is a natural partition of \mathbb{Z} into n parts which generalizes the partition into even and odd. Namely, we put

$$Y_1 = \{\dots, -2n, -n, 0, n, 2n, \dots\} = \{kn \mid k \in \mathbb{Z}\}$$

the set of all multiples of n ,

$$Y_2 = \{\dots, -2n+1, -n+1, 1, n+1, 2n+1, \dots\} = \{kn+1 \mid k \in \mathbb{Z}\},$$

and similarly, for any $0 \leq d \leq n-1$, we put

$$Y_d = \{\dots, -2n+d, -n+d, d, n+d, 2n+d, \dots\} = \{kn+d \mid k \in \mathbb{Z}\}.$$

That is, Y_d is the set of all integers which, upon division by n , leave a remainder of d . Earlier we showed that the remainder upon division by n is a well-defined integer in the range $0 \leq d < n$. Here by “well-defined”, I mean that for $0 \leq d_1 \neq d_2 < n$, the sets Y_{d_1} and Y_{d_2} are disjoint. Recall why this is true: if not, there exist k_1, k_2 such that $k_1n + d_1 = k_2n + d_2$, so $d_1 - d_2 = (k_2 - k_1)n$, so $d_1 - d_2$ is a multiple of n . But $-n < d_1 - d_2 < n$, so the only multiple of n it could possibly be is 0, i.e., $d_1 = d_2$. It is clear that each Y_d is nonempty and that their union is all of \mathbb{Z} , so $\{Y_d\}_{d=0}^{n-1}$ gives a partition of \mathbb{Z} .

The corresponding equivalence relation is called **congruence modulo n** , and written as follows:

$$x \equiv y \pmod{n}.$$

What this means is that x and y leave the same remainder upon division by n .

Proposition 1.19. For integers x, y , the following are equivalent:

- (i) $x \equiv y \pmod{n}$.
- (ii) $n \mid x - y$.

Proof. Suppose that $x \equiv y \pmod{n}$. Then they leave the same remainder, say d , upon division by n : there exist $k_1, k_2 \in \mathbb{Z}$ such that $x = k_1n + d$, $y = k_2n + d$, so $x - y = (k_1 - k_2)n$ and indeed $n \mid x - y$. Conversely, suppose that $x = k_1n + d_1$, $y = k_2n + d_2$, with d_1 and d_2 distinct integers both in the interval $[0, n-1]$. Then, if n divides $x - y = (k_1 - k_2)n + (d_1 - d_2)$, then it also divides $d_1 - d_2$, which as above is impossible since $-n < d_1 - d_2 < n$. \square

Example 1.20. (Fibers of a function) Let $f : X \rightarrow Y$ be a function. We define a relation R on X by $(x_1, x_2) \in R$ iff $f(x_1) = f(x_2)$. This is an equivalence relation. The equivalence class of $[x]$ is called the **fiber over $f(x)$** .

1.7. Extra: composition of relations.

Suppose we have a relation $R \subset X \times Y$ and a relation $S \subset Y \times Z$. We can define a **composite relation** $S \circ R \subset X \times Z$ in a way which will generalize compositions of functions. Compared to composition of functions, composition of relations is much less well-known, although as with many abstract concepts, once it is pointed out to you, you begin to see it “in nature”. This section is certainly optional reading.

The definition is simply this:

$$S \circ R = \{(x, z) \in X \times Z \mid \exists y \in Y \text{ such that } (x, y) \in R \text{ and } (y, z) \in S\}.$$

In other words, we say that x in the first set X relates to z in the third set Z if there exists at least one intermediate element y in the second set such that x relates

to y and y relates to z .

In particular, we can always compose relations on a single set X . As a special case, given a relation R , we can compose it with itself: say

$$R^{(2)} = R \circ R = \{(x, z) \in X \times X \mid \exists y \in X \text{ such that } xRy \text{ and } yRz\}.$$

Proposition 1.21. *For a relation R on X , the following are equivalent:*

- (i) R is transitive.
- (ii) $R^{(2)} \subseteq R$.

Exercise 1.8. *Show that the composition of relations is associative.*

Exercise 1.9. *Show: $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$.*

Exercise 1.10. *Let $X = \{1, \dots, N\}$. To a relation R on X we associate its **adjacency matrix** $M = M(R)$: if $(i, j) \in R$, we put $M(i, j) = 1$; otherwise we put $M(i, j) = 0$. Show that the adjacency matrix of the composite relation R^2 is the product matrix $M(R) \cdot M(R)$ in the sense of linear algebra.*

2. FUNCTIONS

Let X and Y be sets. A **function** $f : X \rightarrow Y$ is a special kind of relation between X and Y . Namely, it is a relation $R \subset X \times Y$ satisfying the following condition: for all $x \in X$ there exists exactly one $y \in Y$ such that $(x, y) \in R$. Because element of y attached to a given element x of X is unique, we may denote it by $f(x)$.

Geometrically, a function is a relation which passes the **vertical line test**: every vertical line $x = c$ intersects the graph of the function in exactly one point. In particular, the domain of any function is all of X .

Example 2.1. *The equality relation $\{(x, x) \mid x \in X\}$ on X is a function: $f(x) = x$ for all x . We call this the **identity function** and denote it by 1_X .*

Example 2.2. *a) Let Y be a set. Then $\emptyset \times Y = \emptyset$, so there is a unique relation on $\emptyset \times Y$. This relation is – vacuously – a function.*

b) Let X be a set. Then $X \times \emptyset = \emptyset$, so there is a unique relation on $X \times \emptyset$, with domain \emptyset . If $X = \emptyset$, then we get the empty function $f : \emptyset \rightarrow \emptyset$. If $X \neq \emptyset$ then the domain is not all of X so we do not get a function.

If $f : X \rightarrow Y$ is a function, the second set Y is called the **codomain** of f . Note the asymmetry in the definition of a function: although every element x of the domain X is required to be associated to a unique element y of Y , the same is not required of elements y of the codomain: there may be multiple elements x in X such that $f(x) = y$, or there may be none at all.

The **image** of $f : X \rightarrow Y$ is $\{y \in Y \text{ such that } y = f(x) \text{ for some } x \in X\}$ ⁴

In calculus one discusses functions with domain some subset of \mathbb{R} and codomain \mathbb{R} . Moreover in calculus a function is usually (but not always...) given by some relatively simple algebraic/analytic expression, and the convention is that the domain is the largest subset of \mathbb{R} on which the given expression makes sense.

⁴Some people call this the **range**, but also some people call the set Y (what we called the codomain) the range, so the term is ambiguous and perhaps best avoided.

Example 2.3.

- a) The function $y = 3x$ is a function from \mathbb{R} to \mathbb{R} . Its range is all of \mathbb{R} .
- b) The function $y = x^2$ is a function from \mathbb{R} to \mathbb{R} . Its range is $[0, \infty)$.
- c) The function $y = x^3$ is a function from \mathbb{R} to \mathbb{R} . Its range is all of \mathbb{R} .
- d) The function $y = \sqrt{x}$ is a function from $[0, \infty)$ to \mathbb{R} . Its range is $[0, \infty)$.
- e) The arctangent $y = \arctan x$ is a function from \mathbb{R} to \mathbb{R} . Its range is $(-\frac{\pi}{2}, \frac{\pi}{2})$.

2.1. The set of all functions from X to Y .

Let X and Y be sets. We denote the set of all functions $f : X \rightarrow Y$ by Y^X . Why such a strange notation? The following simple and useful result gives the motivation. Recall that for $n \in \mathbb{Z}^+$, we put $[n] = \{1, 2, \dots, n\}$, and we also put $[0] = \emptyset$. Thus $\# [n] = n$ for all $n \in \mathbb{N}$.

Proposition 2.4. *Let $m, n \in \mathbb{N}$. Then we have*

$$\# [m]^{[n]} = m^n.$$

In words: the set of all functions from $\{1, \dots, n\}$ to $\{1, \dots, m\}$ has cardinality m^n .

Proof. To define a function $f : \{1, \dots, n\} \rightarrow \{1, \dots, m\}$, we must specify a sequence of elements $f(1), \dots, f(n)$ in $\{1, \dots, m\}$. There are m possible choices for $f(1)$, also m possible choices for $f(2)$, and so forth, up to m possible choices for $f(n)$, and these choices are independent. Thus we have $m \cdots m$ n times $= m^n$ choices overall. \square

2.2. Injective functions.

From the perspective of our course, the most important material on functions are the concepts *injectivity*, *surjectivity* and *bijection* and the relation of these properties with the existence of inverse functions.

A function $f : X \rightarrow Y$ is **injective** if every element y of the codomain is associated to at most one element $x \in X$. That is, f is injective if for all $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$.

Let us meditate a bit on the property of injectivity. One way to think about it is via a horizontal line test: a function is injective if and only if each horizontal line $y = c$ intersects the graph of f in **at most** one point. Another way to think about an injective function is as a function which entails no loss of information. That is, for an injective function, if your friend tells you $x \in X$ and you tell me $f(x) \in Y$, then I can, in principle, figure out what x is because it is uniquely determined.

Consider for instance the two functions $f(x) = x^2$ and $f(x) = x^3$. The first function $f(x) = x^2$ is not injective: if y is any positive real number then there are two x -values such that $f(x) = y$, $x = \sqrt{y}$ and $x = -\sqrt{y}$. Or, in other words, if $f(x) = x^2$ and I tell you that $f(x) = 1$, then you are in doubt as to what x is: it could be either $+1$ or -1 . On the other hand, $f(x) = x^3$ is injective, so if I tell you that $f(x) = x^3 = 1$, then we can conclude that $x = 1$.

How can we verify in practice that a function is injective? One way is to construct an inverse function, which we will discuss further later. But in the special case when $f : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function, the methods of calculus give useful criteria for injectivity.

Before stating the result, let us first recall the definitions of increasing and decreasing functions. A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is **(strictly) increasing** if for all $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2 \implies f(x_1) < f(x_2)$. Similarly, f is **(strictly) decreasing** if for all $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2 \implies f(x_1) > f(x_2)$. Notice that a function which is increasing or decreasing is injective. The “problem” is that a function need not be either increasing or decreasing, although “well-behaved” functions of the sort one encounters in calculus have the property that their domain can be broken up into intervals on which the function is either increasing or decreasing. For instance, the function $f(x) = x^2$ is decreasing on $(-\infty, 0)$ and increasing on $(0, \infty)$.

Theorem 2.5. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function.*

- a) If f is injective, then f is either increasing or decreasing.*
- b) If f is differentiable and either $f'(x) > 0$ for all $x \in \mathbb{R}$ or $f'(x) < 0$ for all $x \in \mathbb{R}$, then f is injective.*

It is something of a sad reflection on our calculus curriculum that useful and basic facts like this are not established in a standard calculus course. However, the full details are somewhat intricate. We sketch a proof below.

Proof. We prove part a) by contraposition: that is, we assume that f is continuous and *neither* increasing nor decreasing, and we wish to show that it is not injective. Since f is not decreasing, there exist $x_1 < x_2$ such that $f(x_1) \leq f(x_2)$. Since f is not increasing, there exist $x_3 < x_4$ such that $f(x_3) \geq f(x_4)$. If $f(x_3) = f(x_4)$. We claim that it follows that there exist $a < b < c$ such that either

Case 1: $f(b) \geq f(a)$ and $f(b) \geq f(c)$, or

Case 2: $f(b) \leq f(a)$ and $f(b) \leq f(c)$.

This follows from a somewhat tedious consideration of cases as to in which order the four points x_1, x_2, x_3, x_4 occur, which we omit here. Now we apply the Intermediate Value Theorem to f on the intervals $[a, b]$ and $[b, c]$. In Case 1, every number smaller than $f(b)$ but sufficiently close to it is assumed both on the interval $[a, b]$ and again on the interval $[b, c]$, so f is not injective. In Case 2, every number larger than $f(b)$ but sufficiently close to it is assumed both on the interval $[a, b]$ and again on $[b, c]$, so again f is not injective.

As for part b), we again go by contraposition and assume that f is not injective: that is, we suppose that there exist $a < b$ such that $f(a) = f(b)$. Applying the Mean Value Theorem to f on $[a, b]$, we get that there exists c , $a < c < b$, such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} = 0,$$

contradicting the assumption that $f'(x)$ is always positive or always negative. \square

Remark: The proof shows that we could have replaced part b) with the apparently weaker hypothesis that for all $x \in \mathbb{R}$, $f'(x) \neq 0$. However, it can be shown that this is equivalent to f' always being positive or always being negative, a consequence of the **Intermediate Value Theorem For Derivatives**.

Example 2.6. a) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = \arctan x$. We claim f is injective. Indeed, it is differentiable and its derivative is $f'(x) = \frac{1}{1+x^2} > 0$ for all $x \in \mathbb{R}$. Therefore f is strictly increasing, hence injective.

b) Let $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = -x^3 - x$. We claim f is injective. Indeed, it is

differentiable and its derivative is $f'(x) = -3x^2 - 1 = -(3x^2 + 1) < 0$ for all $x \in \mathbb{R}$. Therefore f is strictly decreasing, hence injective.

Example 2.7. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $f(x) = x^3$. One meets this function in precalculus and calculus mathematics, and one certainly expects it to be injective. Unfortunately the criterion of Theorem 2.5 falls a bit short here: the derivative is $f'(x) = 3x^2$, which is always non-negative but is 0 at $x = 0$.

We will show “by hand” that f is indeed injective. Namely, let $x_1, x_2 \in \mathbb{R}$ and suppose $x_1^3 = x_2^3$. Then

$$0 = x_1^3 - x_2^3 = (x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2).$$

Seeking a contradiction, we suppose that $x_1 \neq x_2$. Then $x_1 - x_2 \neq 0$, so we can divide through by it, getting

$$0 = x_1^2 + x_1x_2 + x_2^2 = \left(x_1 + \frac{x_2}{2}\right)^2 + \frac{3}{4}x_2^2.$$

Because each of the two terms in the sum is always non-negative, the only way the sum can be zero is if

$$\left(x_1 + \frac{x_2}{2}\right)^2 = \frac{3}{4}x_2^2 = 0.$$

The second equality implies $x_2 = 0$, and plugging this into the first inequality gives $x_1^2 = 0$ and thus $x_1 = 0$. So $x_1 = 0 = x_2$: contradiction.

We gave a proof of the injectivity of $f : x \mapsto x^3$ to nail down the fact that Theorem 2.5 gives a sufficient but not necessary criterion for a differentiable function to be injective. But we would really like to be able to *improve* Theorem 2.5 so as to handle this example via the methods of calculus. For instance, let n be a positive integer. Then we equally well believe that the function $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^{2n+1}$ should be injective. It is possible to show this using the above factorization method....but it is real work to do so. The following criterion comes to the rescue to do this and many other examples easily.

Theorem 2.8. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function.

- a) Suppose that $f'(x) \geq 0$ for all x and that there is no $a < b$ such that $f'(x) = 0$ for all $x \in (a, b)$. Then f is strictly increasing (hence injective).
- b) Suppose that $f'(x) \leq 0$ for all x and that there is no $a < b$ such that $f'(x) = 0$ for all $x \in (a, b)$. Then f is strictly decreasing (hence injective).

Proof. We prove part a); the proof of part b) is identical. Again we go by contrapositive: suppose that f is not strictly increasing, so that there exists $a < b$ such that $f(a) \leq f(b)$. If $f(a) < f(b)$, then applying the Mean Value Theorem, we get a c in between a and b such that $f'(c) < 0$, contradiction. So we may assume that $f(a) = f(b)$. Then, by exactly the same MVT argument, $f'(x) \geq 0$ for all x implies that f is at least weakly increasing, i.e., $x_1 \leq x_2 \implies f(x_1) \leq f(x_2)$. But a weakly increasing function f with $f(a) = f(b)$ must be constant on the entire interval $[a, b]$, hence $f'(x) = 0$ for all x in (a, b) , contradicting the hypothesis. \square

Worked Exercise 2.1. We will show that for any $n \in \mathbb{Z}^+$, the function $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $x \mapsto x^{2n+1}$ is injective. Indeed we have $f'(x) = (2n+1)x^{2n}$, which is non-negative for all $x \in \mathbb{R}$ and is 0 only at $x = 0$. So Theorem 2.8a) applies to show that f is strictly increasing, hence injective.

2.3. Surjective functions. A function $f : X \rightarrow Y$ is surjective if its image $f(X)$ is equal to the codomain Y . More plainly, for all $y \in Y$, there is $x \in X$ such that $f(x) = y$.

In many ways surjectivity is the “dual property” to injectivity. For instance, it can also be verified by a horizontal line test: a function f is surjective if and only if each horizontal line $y = c$ intersects the graph of f in **at least one point**.

Worked Exercise 2.2. Let m and b be real numbers. Is $f(x) = mx + b$ surjective?

Solution: It is surjective if and only if $m \neq 0$. First, if $m = 0$, then $f(x) = b$ is a constant function: it maps all of \mathbb{R} to the single point b and therefore is at the opposite extreme from being surjective. Conversely, if $m \neq 0$, write $y = mx + b$ and solve for x : $x = \frac{y-b}{m}$. Note that this argument also shows that if $m \neq 0$, f is injective: given an arbitrary y , we have solved for a unique value of x .

By the intermediate value theorem, if a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ takes on two values $m \leq M$, then it also takes on every value in between. In particular, if a continuous function takes on arbitrarily large values and arbitrarily small values, then it is surjective.

Theorem 2.9. Let $a_0, \dots, a_n \in \mathbb{R}$ and suppose $a_n \neq 0$. Let $P : \mathbb{R} \rightarrow \mathbb{R}$ by

$$P(x) = a_n x^n + \dots + a_1 x + a_0.$$

Thus P is a polynomial of degree n . Then: P is surjective if and only if n is odd.

Proof. Suppose that n is odd. Then, if the leading term a_n is positive, then

$$\lim_{x \rightarrow \infty} P(x) = +\infty, \quad \lim_{x \rightarrow -\infty} P(x) = -\infty,$$

whereas if the leading term a_n is negative, then

$$\lim_{x \rightarrow \infty} P(x) = -\infty, \quad \lim_{x \rightarrow -\infty} P(x) = +\infty,$$

so either way P takes on arbitrarily large and small values. By the Intermediate Value Theorem, its range must be all of \mathbb{R} .

Now suppose n is even. Then if a_n is positive, we have

$$\lim_{x \rightarrow \infty} P(x) = \lim_{x \rightarrow -\infty} P(x) = +\infty.$$

It follows that there exists a non-negative real number M such that if $|x| \geq M$, $P(x) \geq 0$. On the other hand, since the restriction of P to $[-M, M]$ is a continuous function on a closed interval, it is bounded below: there exists a real number m such that $P(x) \geq m$ for all $x \in [-M, M]$. Therefore $P(x) \geq m$ for all x , so it is not surjective. Similarly, if a_n is negative, we can show that P is bounded above so is not surjective. \square

2.4. Bijective functions.

A function $f : X \rightarrow Y$ is **bijective** if it is both injective and surjective.

Exercise 2.3. Show: on any set X , the identity function $1_X : X \rightarrow X$ by $1_X(x) = x$ is bijective.

Exercise 2.4. Determine which of the functions introduced so far in this section are bijective.

A function is bijective iff for every $y \in Y$, there exists a unique $x \in X$ such that $f(x) = y$.

The following result is easy but of the highest level of importance.

Theorem 2.10. *For a function $f : X \rightarrow Y$, the following are equivalent:*

- (i) *f is bijective.*
- (ii) *The inverse relation $f^{-1} : Y \rightarrow X = \{(f(x), x) \mid x \in X\}$ is itself a function.*

Proof. Indeed, we need f to be surjective so that the domain of f^{-1} is all of Y and we need it to be injective so that each y in Y is associated to no more than one x value. \square

2.5. Composition of functions.

Probably the most important and general property of functions is that they can, under the right circumstances, be *composed*.⁵ For instance, in calculus, complicated functions are built up out of simple functions by plugging one function into another, e.g. $\sqrt{x^2 + 1}$, or $e^{\sin x}$, and the most important differentiation rule – the Chain Rule – tells how to find the derivative of a composition of two functions in terms of the derivatives of the original functions.

Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$: that is, the codomain of f is equal to the domain of g . Then we can define a new function $g \circ f : X \rightarrow Z$ by:

$$x \mapsto g(f(x)).$$

Remark: Note that $g \circ f$ means first perform f and then perform g . Thus function composition proceeds from right to left, counterintuitively at first. There was a time when this bothered mathematicians enough to suggest writing functions *on the right*, i.e., $(x)f$ rather than $f(x)$. But that time is past.

Remark: The condition for composition can be somewhat relaxed: it is not necessary for the domain of g to equal the codomain of f . What is precisely necessary and sufficient is that for every $x \in X$, $f(x)$ lies in the domain of g , i.e.,

$$\text{Range}(f) \subseteq \text{Codomain}(g).$$

Example: The composition of functions is generally *not* commutative. In fact, if $g \circ f$ is defined, $f \circ g$ need not be defined at all. For instance, suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is the function which takes every rational number to 1 and every irrational number to 0 and $g : \{0, 1\} \rightarrow \{a, b\}$ is the function $0 \mapsto b$, $1 \mapsto a$. Then $g \circ f : \mathbb{R} \rightarrow \{a, b\}$ is defined: it takes every rational number to a and every irrational number to b . But $f \circ g$ makes no sense at all:

$$f(g(0)) = f(b) = ???.$$

Remark: Those who have taken linear algebra will notice the analogy with the multiplication of matrices: if A is an $m \times n$ matrix and B is an $n \times p$ matrix, then the product AB is defined, an $m \times p$ matrix. But if $m \neq p$, the product BA is not defined. (In fact this is more than an analogy, since an $m \times n$ matrix A can be viewed as a linear transformation $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Matrix multiplication is indeed

⁵This is a special case of the composition of relations described in §X.X, but since that was optional material, we proceed without assuming any knowledge of that material.

a special case of composition of functions.)

Even when $g \circ f$ and $f \circ g$ are both defined – e.g. when $f, g : \mathbb{R} \rightarrow \mathbb{R}$, they need not be equal. This is again familiar from precalculus mathematics. If $f(x) = x^2$ and $g(x) = x + 1$, then

$$g(f(x)) = x^2 + 1, \text{ whereas } f(g(x)) = (x + 1)^2 = x^2 + 2x + 1.$$

On the other hand, function composition is always **associative**: if $f : X \rightarrow Y$, $g : Y \rightarrow Z$ and $h : Z \rightarrow W$ are functions, then we have

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Indeed the proof is trivial, since both sides map $x \in X$ to $h(g(f(x)))$.⁶

Exercise: Let $f : X \rightarrow Y$.

- a) Show that $f \circ 1_X = f$.
- b) Show that $1_Y \circ f = f$.

2.6. Basic facts about injectivity, surjectivity and composition.

Here we establish a small number of very important facts about how injectivity, surjectivity and bijectivity behave with respect to function composition. First:

Theorem 2.11. *Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be two functions.*

- a) *If f and g are injective, then so is $g \circ f$.*
- b) *If f and g are surjective, then so is $g \circ f$.*
- c) *If f and g are bijective, then so is $g \circ f$.*

Proof. a) We must show that for all $x_1, x_2 \in X$, if $g(f(x_1)) = g(f(x_2))$, then $x_1 = x_2$. But put $y_1 = f(x_1)$ and $y_2 = f(x_2)$. Then $g(y_1) = g(y_2)$. Since g is assumed to be injective, this implies $y_1 = y_2 = f(x_2)$. Since f is also assumed to be injective, this implies $x_1 = x_2$.

b) We must show that for all $z \in Z$, there exists at least one x in X such that $g(f(x)) = z$. Since $g : Y \rightarrow Z$ is surjective, there exists $y \in Y$ such that $g(y) = z$. Since $f : X \rightarrow Y$ is surjective, there exists $x \in X$ such that $f(x) = y$. Then $g(f(x)) = g(y) = z$.

c) Finally, if f and g are bijective, then f and g are both injective, so by part a) $g \circ f$ is injective. Similarly, f and g are both surjective, so by part b) $g \circ f$ is surjective. Thus $g \circ f$ is injective and surjective, i.e., bijective, qed. \square

Now we wish to explore the other direction: suppose we know that $g \circ f$ is injective, surjective or bijective? What can we conclude about the “factor” functions f and g ?

The following example shows that we need to be careful.

Example: Let $X = Z = \{0\}$, let $Y = \mathbb{R}$. Define $f : X \rightarrow Y$ be $f(0) = \pi$ (or your favorite real number; it would not change the outcome), and let f be the constant function which takes every real number y to 0: note that this is the unique function from \mathbb{R} to $\{0\}$. We compute $g \circ f$: $g(f(0)) = g(\pi) = 0$. Thus $g \circ f$ is the identity function on X : in particular it is bijective. However, both f and g are far

⁶As above, this provides a conceptual reason behind the associativity of matrix multiplication.

from being bijective: the range of f is only a single point $\{\pi\}$, so f is not surjective, whereas g maps every real number to 0, so is not injective.

On the other hand, something is true: namely the “inside function” f is injective, and the outside function g is surjective. This is in fact a general phenomenon.

Theorem 2.12. (*Green and Brown Fact*) Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions.

- a) If $g \circ f$ is injective, then f is injective.
- b) If $g \circ f$ is surjective, then g is surjective.
- c) If $g \circ f$ is bijective, then f is injective and g is surjective.

Proof. a) We proceed by contraposition: suppose that f is not injective: then there exist $x_1 \neq x_2$ in X such that $f(x_1) = f(x_2)$. But then $g(f(x_1)) = g(f(x_2))$, so that the distinct points x_1 and x_2 become equal under $g \circ f$: that is, $g \circ f$ is not injective.

b) Again by contraposition: suppose that g is not surjective: then there exists $z \in Z$ such that for no y in Y do we have $z = g(y)$. But then we certainly cannot have an $x \in X$ such that $z = g(f(x))$, because if so taking $y = f(x)$ shows that z is in the range of g , contradiction.

c) If $g \circ f$ is bijective, it is injective and surjective, so we apply parts a) and b). \square

Remark: The name of Theorem 2.12 comes from the Spring 2009 version of Math 3200, when I presented this result using green and brown chalk, decided it was important enough to have a name, and was completely lacking in inspiration.

2.7. Inverse Functions.

Finally we come to the last piece of the puzzle: let $f : X \rightarrow Y$ be a function. We know that the inverse relation f^{-1} is a function if and only if f is injective and surjective. But there is another (very important) necessary and sufficient condition for invertibility in terms of function composition. Before stating it, recall that for a set X , the identity function 1_X is the function from X to X such that $1_X(x) = x$ for all $x \in X$. (Similarly $1_Y(y) = y$ for all $y \in Y$.)

We say that a function $g : Y \rightarrow X$ is the **inverse function** to $f : X \rightarrow Y$ if both of the following hold:

- (IF1) $g \circ f = 1_X$: i.e., for all $x \in X$, $g(f(x)) = x$.
- (IF2) $f \circ g = 1_Y$: i.e., for all $y \in Y$, $f(g(y)) = y$.

In other words, g is the inverse function to f if applying one function and then the other – in either order! – brings us back where we started.

The point here is that g is supposed to be related to f^{-1} , the inverse relation. Here is the precise result:

Theorem 2.13. Let $f : X \rightarrow Y$.

- a) The following are equivalent:
 - (i) f is bijective.
 - (ii) The inverse relation $f^{-1} : Y \rightarrow X$ is a function.

(iii) f has an inverse function g .

b) When the equivalent conditions of part a) hold, then the inverse function g is uniquely determined: it is the function f^{-1} .

Proof. a) We already know the equivalence of (i) and (ii): this is Theorem 2.10 above.

(ii) \implies (iii): Assume (ii), i.e., that the inverse relation f^{-1} is a function. We claim that it is then the inverse function to f in the sense that $f^{-1} \circ f = 1_X$ and $f \circ f^{-1} = 1_Y$. We just do it: for $x \in X$, $f^{-1}(f(x))$ is the unique element of X which gets mapped under f to $f(x)$: since x is such an element and the uniqueness is assumed, we must have $f^{-1}(f(x)) = x$. Similarly, for $y \in Y$, $f^{-1}(y)$ is the unique element x of X such that $f(x) = y$, so $f(f^{-1}(y)) = f(x) = y$.

(iii) \implies (i): We have $g \circ f = 1_X$, and the identity function is bijective. By the Green and Brown Fact, this implies that f is injective. Similarly, we have $f \circ g = 1_Y$ is bijective, so by the Green and Brown Fact, this implies that f is surjective. Therefore f is bijective.⁷

b) Suppose that we have any function $g : Y \rightarrow X$ such that $g \circ f = 1_X$ and $f \circ g = 1_Y$. By the proof of part a), we know that f is bijective and thus the inverse relation f^{-1} is a function such that $f^{-1} \circ f = 1_X$, $f \circ f^{-1} = 1_Y$. Thus

$$g = g \circ 1_Y = g \circ (f \circ f^{-1}) = (g \circ f) \circ f^{-1} = 1_X \circ f^{-1} = f^{-1}.$$

□

In summary, for a function f , being bijective, having the inverse relation (obtained by “reversing all the arrows”) be a function, and having another function g which undoes f by composition in either order, are all equivalent.

⁷A very similar argument shows that g is bijective as well.

Chapter five:

Sets and function

1.1 Section sets

The basic concepts of sets and functions are topics covered in high school math courses and are thus familiar to most university students. We take the intuitive point of view that sets are unordered collections of objects. We first recall some standard terminology and notation associated with sets. When we speak about sets, we usually have a “universal set” U in mind, to which the various sets of our discourse belong.

Definition 1 (Set notation) *A set is an unordered collection of distinct objects. We use the notation $x \in S$ to mean “ x is an element of S ” and $x \notin S$ to mean “ x is not an element of S .” Given two subsets (subcollections) of U , X and Y , we say “ X is a subset of Y ,” written $X \subseteq Y$, if $x \in X$ implies that $x \in Y$. Alternatively, we may say that “ Y is a superset of X .” $X \subseteq Y$ and $Y \supseteq X$ mean the same thing. We say that two subsets X and Y of U are equal if $X \subseteq Y$ and $Y \subseteq X$. We use braces to designate sets when we wish to specify or describe them in terms of their elements: $A = \{a, b, c\}$, $B = \{2, 4, 6, \dots\}$. A set with k elements is called a k -set or set with cardinality k . The cardinality of a set A is denoted by $|A|$.*

Since a set is an unordered collection of distinct objects, the following all describe the same 3-element set

$$\{a, b, c\} = \{b, a, c\} = \{c, b, a\} = \{a, b, b, c, b\}.$$

The first three are simply listing the elements in a different order. The last happens to mention some elements more than once. But, since a set consists of distinct objects, the elements of the set are still just a, b, c . Another way to think of this is:

Two sets A and B are equal if and only if every element of A is an element of B and every element of B is an element of A .

Thus, with $A = \{a, b, c\}$ and $B = \{a, b, b, c, b\}$, we can see that everything in A is in B and everything in B is in A . You might think “When we write a set, the elements are in the order written, so why do you say a set is not ordered?” When we write something down we’re stuck — we have to list them in some order. You can think of a set differently: Write each element on a separate slip of paper and put the slips in a paper bag. No matter how you shake the bag, it’s still the same set.

If we are given that A is a set and no other information about A , then there is no ordering to the elements of A . Thus, we cannot speak of “the second element of the set A ” unless we have specified an ordering of the elements of A . If we wish to regard A as ordered in some way, then we specify this fact explicitly: “The elements of A are ordered a, b, c ,” or “ $A = (a, b, c)$.” The latter notation replaces the braces with parentheses and designates that A is ordered, left to right, as indicated. We call this an *ordered set*. An ordered set is also called a *linear order*. Various other names are also used: *list*, *vector*, *string*, *word*

Sets and Functions

— all with **no repeated elements**.¹ Of course, you’ve seen repeated elements in vectors, for example the point in the plane at the coordinates (1,1). That’s fine, it’s just not an ordered set. If there are k elements in the ordered set, it is referred to as a k -list, k -vector, etc., or as a list, vector, etc., of *length* k — all with no repeated elements because they are ordered **sets**.

Sometimes we cannot list the elements of a set explicitly. What do we do if we want to describe the set of all real numbers greater than 1 without writing it out in words? We write

$$\{x \mid x \in \mathbb{R}, x > 1\} \quad \text{or} \quad \{x \mid x > 1\} \quad \text{or} \quad \{x : x > 1\}.$$

These are read “the set of all x such that . . .” In the first example we mentioned that x was a real number ($x \in \mathbb{R}$). In the other two we didn’t because we assumed the reader knew from context that we were talking about real numbers.

For the most part, we shall be dealing with finite sets. Let U be a set and let A and B be subsets of U . The sets

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}$$

and

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$$

are the *intersection* and *union* of A and B . The set $A \setminus B$ or $A - B$ is the *set difference* of A and B (i.e., the set $\{x \mid x \in A, x \notin B\}$). The set $U \setminus A$ (also A^c , A' or $\sim A$) is the *complement* of A (relative to U). Note that $A - B = \{x \mid x \in A, x \notin B\} = A \cap B^c$. The empty set, denoted by \emptyset , equals U^c . Also note that, for any set $A \subseteq U$, $A \cup A^c = U$ and $A \cap A^c = \emptyset$.

The set $A \oplus B = (A \setminus B) \cup (B \setminus A)$ is the *symmetric difference* of A and B . We use $A \times B = \{(x, y) \mid x \in A, y \in B\}$ to denote the *product* or *Cartesian product* of A and B . If we want to consider the product of k sets, A_1, \dots, A_k , this is denoted by $\times_{i=1}^k A_i$. If we want to consider the product of a set A with itself k times, we write $\times^k A$.

Set Properties and Proofs

The algebraic rules for operating with sets are also familiar to most beginning university students. Here is such a list of the basic rules. In each case the standard name of the rule is given first, followed by the rule as applied first to \cap and then to \cup .

Theorem 1 (Algebraic rules for sets) *The universal set U is not mentioned explicitly but is implicit when we use the notation $\sim X = U - X$ for the complement of X . An*

¹ Why is it okay to specify a set $S = \{a, b, c, a\}$ where the element a has been repeated, but it is not okay to have repeated elements in an ordering of S ? When we say $S = \{a, b, c, a\}$, we know that S contains just the three elements a , b and c . If we were to talk about the ordered set (a, b, c, a) it would not make sense because it would say that the element a is in two places at once: the first position and the last position.

alternative notation is $X^c = \sim X$.

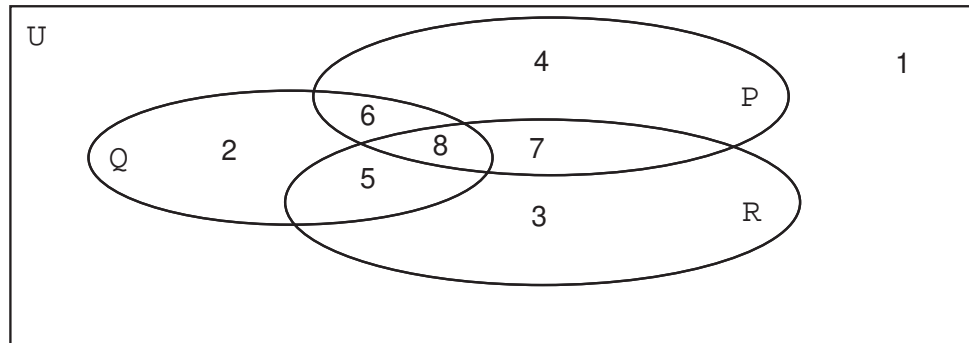
| | | |
|------------------|--|--|
| Associative: | $(P \cap Q) \cap R = P \cap (Q \cap R)$ | $(P \cup Q) \cup R = P \cup (Q \cup R)$ |
| Distributive: | $P \cap (Q \cup R) = (P \cap Q) \cup (P \cap R)$ | $P \cup (Q \cap R) = (P \cup Q) \cap (P \cup R)$ |
| Idempotent: | $P \cap P = P$ | $P \cup P = P$ |
| Double Negation: | $\sim \sim P = P$ | |
| DeMorgan: | $\sim(P \cap Q) = \sim P \cup \sim Q$ | $\sim(P \cup Q) = \sim P \cap \sim Q$ |
| Absorption: | $P \cup (P \cap Q) = P$ | $P \cap (P \cup Q) = P$ |
| Commutative: | $P \cap Q = Q \cap P$ | $P \cup Q = Q \cup P$ |

These rules are “algebraic” rules for working with \cap , \cup , and \sim . You should memorize them as you use them. They are used just like rules in ordinary algebra: whenever you see an expression on one side of the equal sign, you can replace it by the expression on the other side.

When we wrote “ $P \cap Q \cap R$ ” you may have wondered if we meant “ $(P \cap Q) \cap R$ ” or “ $P \cap (Q \cap R)$.” The associative law says it doesn’t matter. That is why you will see the notation $P \cap Q \cap R$ or $P \cup Q \cup R$ without anyone getting excited about it. On the other hand $P \cap (Q \cup R)$ and $(P \cap Q) \cup R$ may not be equal, so we need parentheses here.

The best way to “prove” the rules or to understand their validity is through the geometric device of a *Venn diagram*.

Example 1 (Venn diagrams and proofs of set equations) Here is a Venn diagram for three sets, P , Q , and R , with universal set U :



The three oval regions labeled P , Q , and R represent the sets of those names. The rectangular region represents the universal set U . There are eight subregions, labeled 1 through 8 in the picture. Region 8 represents the subset $P \cap Q \cap R$; region 1 represents $U - (P \cup Q \cup R)$; region 2 represents the elements of $Q - (P \cup R)$; and so on.

Let’s use the above Venn diagram to verify that the distributive rule, $P \cup (Q \cap R) = (P \cup Q) \cap (P \cup R)$, is valid. The idea is to replace the sets P , Q , and R by their corresponding sets of regions from the Venn diagram. Thus, Q is replaced by $\{2, 5, 6, 8\}$, P is replaced by $\{4, 6, 7, 8\}$, and R is replaced by $\{3, 5, 7, 8\}$. Even though the sets P , Q , and R are arbitrary, perhaps even infinite, the distributive rule reduces to verifying the same rule for these simplified sets:

$$\{4, 6, 7, 8\} \cup (\{2, 5, 6, 8\} \cap \{3, 5, 7, 8\}) = (\{4, 6, 7, 8\} \cup \{2, 5, 6, 8\}) \cap (\{4, 6, 7, 8\} \cup \{3, 5, 7, 8\}).$$

Sets and Functions

This identity is trivial to check directly: Both sides reduce to the set $\{4, 5, 6, 7, 8\}$.

This “Venn diagram” approach reduces a set identity that involves potentially infinitely many elements to subsets of a set of eight elements. It is fine for proofs and especially good for checking out “set identities” to see quickly if they are true or not. For example, is it true that $Q - (P \cap R) = Q - (P \cap Q \cap R)$? Checking the Venn diagram shows that both sides correspond to the set of regions $\{2, 5, 6\}$. The identity is true. You will get a chance to practice this technique in the exercises. \square

There are, of course, other ways to verify set identities. One way is called the *element method*:

Example 2 (The element method for proofs of set equations) To use that method, you simply translate the identity $X = Y$ into basic statements about what conditions a single element must satisfy to be (first) in the set on the left and then (second) in the set on the right. Thus, to show that $X = Y$, you assert that if $x \in X$ then blah, blah, blah (a bunch of words that make sense) implies that $x \in Y$. This shows that $X \subseteq Y$. Then, you reverse the argument and assert that if $y \in Y$ then blah, blah, blah (a bunch of words that make sense) implies that $y \in X$. This shows that $Y \subseteq X$. Thus $X = Y$.

Here is an example. Show, by the element method that, for all subsets P , Q , and R of U , $(P - Q) \cap (R - Q) = (P \cap R) - Q$.

- (1) If $x \in (P - Q) \cap (R - Q)$ then (here comes the blah, blah, blah) x is in P but not in Q AND x is in R but not in Q .
- (2) Thus x is in P **and** R , but x is not in Q .
- (3) Thus x is in $(P \cap R) - Q$. This shows that $(P - Q) \cap (R - Q) \subseteq (P \cap R) - Q$. We leave it to you to use the element method to show the reverse, $(P - Q) \cap (R - Q) \supseteq (P \cap R) - Q$, and hence that $(P - Q) \cap (R - Q) = (P \cap R) - Q$. You should start your argument by saying, “Suppose $x \in (P \cap R) - Q$.” \square

A different sort of element approach looks at each element of the universal set U and asks which sets contain it. The result can be put in tabular form. When this is done, each row of the table corresponds to a region in the Venn diagram. The next example illustrates this *tabular method*.

Example 3 (The tabular method for proofs of set equations) We redo the identity of the previous example: $(P - Q) \cap (R - Q) = (P \cap R) - Q$. To do this we construct a table whose columns are labeled by various sets and whose entries answer the question “Is x in the set?” The first three columns in the following table are set up to allow all possible answers to the three questions “Is x in P ?” “Is x in Q ?” “Is x in R ?” “Left” and “Right” refer to $(P - Q) \cap (R - Q)$ and $(P \cap R) - Q$, the two sides of the equation we want to prove. “Venn” refers to the region in the Venn diagram of Example 1. Normally that column would not be in the table, but we’ve inserted it so that you can see how each row

corresponds to a Venn diagram region.

| P | Q | R | $P - Q$ | $R - Q$ | Left | $P \cap R$ | Right | Venn |
|-----|-----|-----|---------|---------|------|------------|-------|------|
| No | No | No | No | No | No | No | No | 1 |
| No | No | Yes | No | Yes | No | No | No | 3 |
| No | Yes | No | No | No | No | No | No | 2 |
| No | Yes | Yes | No | No | No | No | No | 5 |
| Yes | No | No | Yes | No | No | No | No | 4 |
| Yes | No | Yes | Yes | Yes | Yes | Yes | Yes | 7 |
| Yes | Yes | No | No | No | No | No | No | 6 |
| Yes | Yes | Yes | No | No | No | Yes | No | 8 |

Since the answers are identical in the columns labeled “Left” and “Right,” the identity is proved.

We can prove more from the table. For example,

$$\text{If } P \subseteq Q \cup R, \text{ then } P - Q = (P \cap R) - Q.$$

How does the table prove this? Because of the condition, the row that begins “Yes No No” is impossible. Therefore, we throw out that row and compare columns “ $P - Q$ ” and “Right.” \square

Another way to prove set identities is to use the basic algebraic identities of Theorem 1. This is called the *algebraic method*.

Example 4 (An algebraic proof) It is probably a good idea for you to label the steps with the appropriate rule (e.g., DeMorgan’s rule, associative rule, distributive rule, etc.) the first few times you do such a proof. Therefore, we’ll do that in this example. Mathematicians, however, would rarely bother to do it. A proof is accepted if others who know the basic rules of set theory can read it, understand it, and believe it is true.

Let’s prove that $Q - (P \cap R) = Q - (P \cap Q \cap R)$. Here it is

$$\begin{aligned}
 Q - (P \cap R) &= Q \cap (P \cap R)^c && \text{since } A - B = A \cap B^c \\
 &= Q \cap (P^c \cup R^c) && \text{DeMorgan’s rule} \\
 &= (Q \cap P^c) \cup (Q \cap R^c) && \text{distributive rule} \\
 &= (Q \cap P^c) \cup \emptyset \cup (Q \cap R^c) && \text{since } A \cup \emptyset = A \\
 &= (Q \cap P^c) \cup (Q \cap Q^c) \cup (Q \cap R^c) && \text{since } Q \cap Q^c = \emptyset \\
 &= Q \cap (P^c \cup Q^c \cup R^c) && \text{distributive rule} \\
 &= Q \cap (P \cap Q \cap R)^c && \text{DeMorgan’s rule} \\
 &= Q - (P \cap Q \cap R) && \text{since } A - B = A \cap B^c
 \end{aligned}$$

Some steps in this proof are baffling. For example, why did we introduce \emptyset in the fourth line? We knew where we were going because we worked from both “ends” of the proof. In other words, we came up with a proof that moved both ends toward the middle and then

Sets and Functions

rearranged the steps so that we could go from one end to the other. Unfortunately, proofs are often presented this way.

Here's another way to write the proof of $Q - (P \cap R) = Q - (P \cap Q \cap R)$ that shows more clearly how we got the proof. Note first that this identity is equivalent to showing that

$$Q \cap (P \cap R)^c = Q \cap (P \cap Q \cap R)^c$$

since $A - B = A \cap B^c$. This is equivalent, by DeMorgan's rules, to showing that

$$Q \cap (P^c \cup R^c) = Q \cap (P^c \cup Q^c \cup R^c).$$

But

$$Q \cap (P^c \cup Q^c \cup R^c) = Q \cap ((P^c \cup R^c) \cup Q^c) = (Q \cap (P^c \cup R^c)) \cup (Q \cap Q^c) = Q \cap (P^c \cup R^c).$$

This latter identity follows from the fact that $Q \cap Q^c = \emptyset$ and $X \cup \emptyset = X$ for any set X . This completes the proof.

How should you write an algebraic proof? You can use whichever method you prefer. The first approach can be read mechanically because of the way it's laid out. However, if you use the first approach, you may sometimes need to use the second method for yourself first. \square

Ordering Sets

In computer programming, you will store and compute with sets of all sorts (sets of number, letters, geometric figures, addresses to arrays, pointers to structures, etc.). In almost all cases, you will work with these sets as lists (also called "linear orders") of some type where order does matter. The order matters in terms of the efficiency of your computations, not in terms of the rules of set theory.

In many cases, the linear ordering of the elements of a set is inherited from the universal set U . For example, the sets $A = \{1, 2, 3\}$ and $B = \{a, b, c\}$ inherit a natural linear ordering from the integers and the alphabet, respectively. But what about $C = \{?, >, <\}$? There is no standard convention for C . You could use the ASCII code order ($<, >, ?$), but if you do, some explanation should be given.

Example 5 (Lexicographic order) If you have decided on linear orders (i.e., listings) for a set X and a set Y , there is a commonly used and natural linear ordering for $X \times Y$ called *lexicographic order*. Suppose we list X and Y in some manner:

$$(x_1, x_2, \dots, x_n) \quad \text{and} \quad (y_1, y_2, \dots, y_m).$$

Given pairs $(a, b) \in X \times Y$ and $(c, d) \in X \times Y$, we say that (a, b) is *lexicographically less than or equal to* (c, d) if

- (1) a is before c in the linear order on X or
- (2) $a = c$ and b is equal to or before d in the linear order on Y .

For example, the lexicographic order for $\{1, 2, 7\} \times \{a, b\}$ is

$$\{(1, a), (1, b), (2, a), (2, b), (7, a), (7, b)\}.$$

Lexicographic order is called *lex order* for short.

Once you have ordered $X \times Y$ lexicographically, you can order $(X \times Y) \times Z$ by the same two rules (1) and (2) above, provided an order is specified on Z . You can use lex order on $X \times Y$ and the given linear order on Z . Likewise, you can apply (1) and (2) to $X \times (Y \times Z)$ using the given linear order on X and lex order on $Y \times Z$. The sets $(X \times Y) \times Z$ and $X \times (Y \times Z)$ are different sets – elements in the former have the form $((x, y), z)$ and those in the latter have the form $(x, (y, z))$. Imagine listing $(X \times Y) \times Z$ in lex order and stripping off the inner parentheses so that $((x, y), z)$ becomes (x, y, z) . Now do the same with $X \times (Y \times Z)$. The two lists will contain the same elements in the same order. It's easy to see why the elements are the same, but it's not so easy to see why the orders are the same. Let's prove the orders are the same.

That means we have to prove

$$((x_1, y_1), z_1) \text{ precedes } ((x_2, y_2), z_2) \text{ if and only if } (x_1, (y_1, z_1)) \text{ precedes } (x_2, (y_2, z_2)).$$

It will make things easier if we write “ $((x_1, y_1), z_1) < ((x_2, y_2), z_2)$ ” for “ $((x_1, y_1), z_1)$ precedes $((x_2, y_2), z_2)$.” By definition the definition of lex order, $((x_1, y_1), z_1) < ((x_2, y_2), z_2)$ means that $(x_1, y_1) < (x_2, y_2)$ or $(x_1, y_1) = (x_2, y_2)$ and $z_1 < z_2$. By the definition of lex order, the first case means that either $x_1 < x_2$ or $x_1 = x_2$ and $y_1 < y_2$. Note that $(x_1, y_1) = (x_2, y_2)$ means $x_1 = x_2$ and $y_1 = y_2$. Putting all this together, we have shown that $((x_1, y_1), z_1) < ((x_2, y_2), z_2)$ means that either

- (1) $x_1 < x_2$ or
- (2) $x_1 = x_2$ and $y_1 < y_2$ or
- (3) $x_1 = x_2$ and $y_1 = y_2$ and $z_1 < z_2$.

In other words, look for the first position where (x_1, y_1, z_1) and (x_2, y_2, z_2) disagree and use that position to determine the order. We leave it to you to show that the same conditions describe $(x_1, (y_1, z_1)) < (x_2, (y_2, z_2))$. This completes the proof.

Since $(X \times Y) \times Z$ and $X \times (Y \times Z)$ have the same order when inner parentheses are dropped, one usually does that. We just write $X \times Y \times Z$ and call the elements (x, y, z) .

Sets and Functions

When you think about what we've just done, you should be able to see that, if S_1, S_2, \dots, S_n are sets, we can write $S_1 \times S_2 \times \dots \times S_n$, leaving out parentheses. The product $S_1 \times S_2 \times \dots \times S_n$ is also written $\times_{i=1}^n S_i$.

Suppose (s_1, s_2, \dots, s_n) and (t_1, t_2, \dots, t_n) are in $S_1 \times S_2 \times \dots \times S_n$. Which comes first? You should be able to see that we determine which precedes which as follows: By going left to right, find the first position where they disagree. Say position k is where this disagreement occurs. Use the order of s_k and t_k to determine the order of (s_1, s_2, \dots, s_n) and (t_1, t_2, \dots, t_n) . This is the same order you use when you look things up in a dictionary. \square

Example 6 (Dictionary order on words or strings) The order of words in the dictionary is called "dictionary order." Lex order appears to be the same as dictionary order, but there is a problem with this. We've only defined lex order for n -tuples where n has some fixed value, but words in the dictionary have different length. Let's look at this more carefully.

Let S be a finite set. Let $S^k = \times^k S$ be the product of S with itself k times. The set S^0 is special and consists of one string ϵ called the *empty string*. Let $S^* = \cup_{k=0}^{\infty} S^k$. In words, S^* consists of all strings (words, vectors) of all possible lengths (including length zero) over S . Assume S is linearly ordered. We now define an order relation on S^* called *lexicographic order* or *dictionary order*, denoted by \leq_L , on S^* .

Let (a_1, a_2, \dots, a_m) and (b_1, b_2, \dots, b_n) be two elements of S^* with $m, n > 0$. We say that

$$(a_1, a_2, \dots, a_m) \leq_L (b_1, b_2, \dots, b_n)$$

if either of the following two conditions hold:

(D1) $m \leq n$ and $a_i = b_i$ for $i = 1, \dots, m$.

(D2) For some $k < \min(m, n)$, $a_i = b_i$, $i = 1, \dots, k$, $a_{k+1} \neq b_{k+1}$, and a_{k+1} is before b_{k+1} in the linear order on S .

Since $m, n > 0$, we have not discussed the empty string. Thus we need:

(D3) The empty string $\epsilon \leq_L x$ for any string x .

We have just defined dictionary order and also called it lex order. Is this the same as our previous definition of lex order? Yes because the two definitions of lex order agree when the strings have the same length.

We shall study this ordering on words carefully when we study order relations in general. For now we just give an example. Let $S = \{x, y\}$ with the ordering on S the alphabetic order. If $u = (x, x, y)$ and $v = (x, x, y, x)$, then $u \leq_L v$ by (D1). If $s = (x, x, y, x)$ and $t = (x, x, x, y)$, then $t \leq_L s$ by (D2). More examples will be given in the exercises. The standard English dictionary is an example where this linear order is applied to a subset of all words on the standard English alphabet (the words that have meaning in English).

A variation on this dictionary order is to order all words first by length and then by lex order. Thus, $u = (y, y, y)$ comes before $v = (x, x, x, x)$ because u has length three (three components) and v has length one. This order on S^* is called *length-first lex order* or *short lex order*. \square

Subsets of Sets

We use the notation $\mathcal{P}(A)$ to denote the set of all subsets of A and $\mathcal{P}_k(A)$ the set of all subsets of A of size (or cardinality) k . We call $\mathcal{P}(A)$ “the set of all subsets of A ” or simply the *power set* of A . Let $C(n, k) = |\mathcal{P}_k(A)|$ denote the number of different k -subsets that can be formed from an n -set. The notation $\binom{n}{k}$ is also frequently used. These are called *binomial coefficients* and are read “ n choose k .” We now prove

Theorem 2 (Binomial coefficient formula) *The value of the binomial coefficient is*

$$\binom{n}{k} = C(n, k) = \frac{n(n-1) \cdots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!},$$

where $0! = 1$ and, for $j > 0$, $j!$ is the product of the first j integers. We read $j!$ as “ j factorial”.

Proof: Let A be a set of size n . The elements of $\mathcal{P}_k(A)$ are sets and are thus unordered. Generally speaking, unordered things are harder to count than ordered ones. Suppose, instead of a set of size k chosen from A , you wanted to construct an ordered list L of k elements from A (L is called a “ k -list”). We could construct L in two stages.

- First choose an element of $S \in \mathcal{P}_k(A)$ (a subset of A with k elements). This can be done in $C(n, k)$ ways since $C(n, k) = |\mathcal{P}_k(A)|$.
- Next order S to obtain L . This ordering can be done in $k! = k(k-1) \cdots 1$ ways. Why? You have k choices for the element of S to appear first in the list L , $k-1$ choices for the next element, $k-2$ choices for the next element, etc.

From this two-stage process, we see that there are $C(n, k) k!$ ordered k -lists with no repeats. (The factor $C(n, k)$ is the number of ways to carry out the first stage and the factor $k!$ is the number of ways to carry out the second stage.)

Theorem 3 (Number of ordered lists) *The number of ordered k -lists L that can be made from an n -set A is*

- n^k if repeats are allowed and
- $n(n-1) \cdots (n-k+1) = n!/(n-k)!$ if repeats are not allowed. One also uses the notation $(n)_k$ for these values. This is called the “falling factorial” and is read “ n falling k ”.

Why? With repeats allowed, there are n choices of elements in A for the first entry in the k -list L , n choices for the second entry, etc. If repeats are not allowed, there are n choices of elements in A for the first entry in the k -list L , $n-1$ choices for the second entry, etc.

Sets and Functions

Since we've counted the same thing (k -lists made from A) in two different ways, the two answers must be equal; that is, $C(n, k)k! = n!/(n - k)!$. Dividing by $k!$, we have the theorem. \square

In high school, you learned about "Pascal's Triangle" for computing binomial coefficients. We review this idea in the next example.

Example 7 (Binomial recursion) Let $X = \{x_1, \dots, x_n\}$. We'll think of $C(n, k)$ as counting k -subsets of X . Imagine that we are going to construct a subset S of X with k elements. Either the element x_n is in our subset S or it is not. The cases where it is in the subset S are all formed by taking the various $(k - 1)$ -subsets of $X - \{x_n\}$ and adding x_n to them. By the definition of binomial coefficients, there are $\binom{n-1}{k-1}$ such subsets. The cases where it is not in the subset S are all formed by taking the various k -subsets of $X - \{x_n\}$. By the definition of binomial coefficients, there are $\binom{n-1}{k}$ such subsets. What we've done is describe how to build all k -subsets of X from certain subsets of $X - \{x_n\}$. Since this gives each subset exactly once,

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k},$$

which can be written $C(n, k) = C(n - 1, k - 1) + C(n - 1, k)$. This equation is called a *recursion* because it tells how to compute the function $C(n, k)$ from values of the function with smaller arguments. Here are the starting values together with the basic recursion:

$$C(1, 0) = C(1, 1) = 1,$$

$$C(1, k) = 0 \quad \text{for } k \neq 0, 1 \quad \text{and}$$

$$C(n, k) = C(n - 1, k - 1) + C(n - 1, k) \quad \text{for } n > 1.$$

Below we have made a table of values for $C(n, k)$.

| $n \backslash k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------|---|---|----|----|----|---|---|
| 0 | 1 | | | | | | |
| 1 | 1 | 1 | | | | | |
| 2 | 1 | 2 | 1 | | | | |
| 3 | 1 | 3 | 3 | 1 | | | |
| 4 | 1 | 4 | 6 | 4 | 1 | | |
| 5 | 1 | 5 | 10 | 10 | 5 | 1 | |
| 6 | 1 | 6 | 15 | 20 | 15 | 6 | 1 |

This tabular representation of $C(n, k)$ is called "Pascal's Triangle." \square

Definition 2 (Characteristic function) Let U be the universal set and let $A \subseteq U$. The characteristic function of A , denoted χ_A is defined for each $x \in U$ by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

Thus the domain of χ_A is U and the range of χ_A is $\{0, 1\}$.²

² If you are not familiar with "domain" and "range", see the definition at the beginning

Example 8 (Subsets as (0,1)-vectors) If A has n elements, listed (a_1, a_2, \dots, a_n) , then you can specify any subset $X \subset A$ by a sequence $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ where $\epsilon_k = 0$ if the element $a_k \notin X$ and $\epsilon_k = 1$ if the element $a_k \in X$. The vector $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ is just the characteristic function of X since $\epsilon_k = \chi_X(a_k)$.

How many different subsets of A are there? We'll show that there are 2^n choices for $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$ and thus $|\mathcal{P}(A)| = 2^n$. Why 2^n ? There are clearly two choices for ϵ_1 and two choices for ϵ_2 and so forth. Thus there are $2 \times 2 \times \dots = 2^n$ choices for $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$. \square

Example 9 (Sets with sets as elements) Sets can have sets as elements. In the first exercise of this section, you will be asked such questions as “Is $\{1, 2\} \in \{\{1, 2\}, \{3, 4\}\}?”$ or “Is $1 \in \{\{1\}, \{2\}, \{3\}\}?”$ Easy stuff if you understand the definitions: You can see that the set $\{1, 2\}$ is indeed an element of the set $\{\{1, 2\}, \{3, 4\}\}$ because this latter set has just two elements, each of them a set of size two, one of which is $\{1, 2\}$. You can also see that every element of $\{\{1\}, \{2\}, \{3\}\}$ is a set and that the number 1 is nowhere to be found as an element of this set.

You have already seen $\mathcal{P}(A)$, which is a set whose elements are sets, namely the subsets of A .

Another important class of sets with sets as elements are the *set partitions*. Some of the elementary aspects of set partitions fit into our present discussion. More advanced aspects of them will be discussed in Section 2. Here is a preview. Let $A = \{1, 2, \dots, 15\}$. Consider the following set whose elements are themselves subsets of A .

$$\alpha = \{\{1\}, \{2\}, \{9\}, \{3, 5\}, \{4, 7\}, \{6, 8, 10, 15\}, \{11, 12, 13, 14\}\}.$$

This set is a subset of the power set $\mathcal{P}(A)$. But, it is a very special type of subset, called a *set partition* of A because it satisfies the three conditions:

- (1) every element of α is nonempty,
- (2) the union of the elements of α is A , and
- (3) if you pick sets $X \in \alpha$ and $Y \in \alpha$, either $X = Y$ or $X \cap Y = \emptyset$.

Any collection of subsets of a set A satisfying (1), (2), and (3) is a set partition of A or simply a partition of A . Each element of α (which is, of course, a subset of A) is called a *block* of the partition α .

How many partitions are there of a set A ? This is a tricky number to compute and there is no simple formula like $C(n, k) = \frac{n!}{k!(n-k)!}$ for it. We will discuss it in the Section 2. The number of partitions of a set of size n is denoted by B_n . These numbers are called *Bell numbers* after Eric Temple Bell. The first few Bell numbers are $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$.

We can *refine* the partition α by splitting blocks into smaller blocks. For example, we might split the block $\{6, 8, 10, 15\}$ into two blocks, say $\{6, 15\}$ and $\{8, 10\}$, and also split the block $\{11, 12, 13, 14\}$ into three blocks, say $\{13\}$, $\{14\}$, and $\{11, 12\}$. The resulting partition is called a *refinement* of α and equals

$$\{\{1\}, \{2\}, \{9\}, \{3, 5\}, \{4, 7\}, \{6, 15\}, \{8, 10\}, \{13\}, \{14\}, \{11, 12\}\}.$$

of the next section.

Sets and Functions

Note that a refinement of a partition is another partition of the same set. We also consider a partition α to be a refinement of itself. We shall gain a deeper understanding of the notion of refinement when we study order relations. \square

Exercises for Section 1

1.1. Answer the following about the \in and \subseteq operators.

- (a) Is $\{1, 2\} \in \{\{1, 2\}, \{3, 4\}\}$?
- (b) Is $\{2\} \in \{1, 2, 3, 4\}$?
- (c) Is $\{3\} \in \{1, \{2\}, \{3\}\}$?
- (d) Is $\{1, 2\} \subseteq \{1, 2, \{1, 2\}, \{3, 4\}\}$?
- (e) Is $1 \in \{\{1\}, \{2\}, \{3\}\}$?
- (f) Is $\{1, 2, 1\} \subseteq \{1, 2\}$?

1.2. For each of the following, draw a Venn diagram.

- (a) $A \subseteq B$, $C \subseteq B$, $A \cap C = \emptyset$
- (b) $A \supseteq C$, $B \cap C = \emptyset$.

1.3. Let $A = \{w, x, y, z\}$ and $B = \{a, b\}$. Take the linear orders on A and B to be alphabetic order. List the elements in each of the following sets in lexicographic order.

- (a) $A \times B$
- (b) $B \times A$
- (c) $A \times A$
- (d) $B \times B$

1.4. Let $A = \{1, 2, 3\}$, $B = \{u, v\}$, and $C = \{m, n\}$. Take the linear order on A to be numeric and the linear orders on B and C to be alphabetic. List the elements in each of the following sets in lexicographic order.

- (a) $A \times (B \times C)$ (use lex order on $B \times C$).
- (b) $(A \times B) \times C$ (use lex order on $A \times B$).
- (c) $A \times B \times C$.

1.5. Let $\Sigma = (x, y)$ be an alphabet. List each of the following sets of strings over this alphabet in the order indicated.

Section 1: Sets

- (a) All palindromes (strings that read the same forward and backward) of length less than or equal to 4. List them in dictionary order.
 - (b) All strings (words) that begin with x and have length less than four. List them in both dictionary and length-first lex order.
 - (c) List all strings of length four in lex order.
- 1.6.** Each of the following statements about subsets of a set U is FALSE. Draw a Venn diagram to represent the situation being described. In each case case, show that the assertion is false by specializing the sets.
- (a) For all A , B , and C , if $A \not\subseteq B$ and $B \not\subseteq C$ then $A \not\subseteq C$.
 - (b) For all sets A , B , and C , $(A \cup B) \cap C = A \cup (B \cap C)$.
 - (c) For all sets A , B , and C , $(A - B) \cap (C - B) = A - (B \cup C)$.
 - (d) For all A , B , and C , if $A \cap C \subseteq B \cap C$ and $A \cup C \subseteq B \cup C$ then $A = B$.
 - (e) For all A , B , and C , if $A \cup C = B \cup C$ then $A = B$.
 - (f) For all sets A , B , and C , $(A - B) - C = A - (B - C)$.
- 1.7.** Prove each statement directly from the definitions.
- (a) If A , B , and C are subsets of U , then $A \subseteq B$ and $A \subseteq C$ implies that $A \subseteq B \cap C$.
 - (b) If A , B , and C are subsets of U , then $A \subseteq B$ and $A \subseteq C$ implies that $A \subseteq B \cup C$.
- 1.8.** Prove, using the definition of set equality, that for all sets A , B , and C , $(A - B) \cap (C - B) = (A \cap C) - B$.
- 1.9.** Prove each statement by the method indicated.
- (a) Prove using element arguments that if U is the universal set and A and B subsets of U , then $A \subseteq B$ implies that $U - A \supseteq U - B$ (alternative notation: $A \subseteq B$ implies $A^c \supseteq B^c$, or $A' \supseteq B'$)
 - (b) Prove, using element arguments and the definition of set inclusion, that for all A , B , and C , if $A \subseteq B$ then $A \cap C \subseteq B \cap C$.
 - (c) Prove, using (a), (b), and DeMorgan's law, that for all A , B , and C , if $A \subseteq B$ then $A \cup C \subseteq B \cup C$.
- 1.10.** Prove each statement by the "element method."
- (a) If A , B , and C are subsets of U , then $A \times (B \cup C) = (A \times B) \cup (A \times C)$.
 - (b) If A , B , and C are subsets of U , then $A \times (B \cap C) = (A \times B) \cap (A \times C)$.
- 1.11.** Prove each of the following identities from the basic algebraic rules for sets. You may want to use the fact that $D - E = D \cap E^c$.

Sets and Functions

- (a) If A , B , and C are subsets of U , then $(A - B) - C = A - (B \cup C)$.
 - (b) If A , B , and C are subsets of U , then $(A - B) - C = (A - C) - B$.
 - (c) If A and B are subsets of U , then $(A - B) \cup (B - A) = (A \cup B) - (A \cap B)$.
- 1.12.** Prove or give a counterexample. Use a Venn diagram argument for the proof. For the counterexample, use a Venn diagram or use set specialization.
- (a) If A , B , and C are subsets of U , then $(A - C) \cap (B - C) \cap (A - B) = \emptyset$.
 - (b) If A and B are subsets of U and if $A \subseteq B$, then $A \cap (U - B) = \emptyset$.
 - (c) If A , B , and C are subsets of U , and if $A \subseteq B$, then $A \cap (U - (B \cap C)) = \emptyset$.
 - (d) If A , B , and C are subsets of U , and if $(B \cap C) \subseteq A$, then $(A - B) \cap (A - C) = \emptyset$.
 - (e) If A and B are subsets of U and if $A \cap B = \emptyset$, then $A \times B = \emptyset$.
- 1.13.** Recall that the symmetric difference of sets A and B is $A \oplus B = (A - B) \cup (B - A)$. It is evident from the definition that $A \oplus B = B \oplus A$, the commutative law. Let U be the universal set. Prove each of the following properties either using a Venn diagram argument or algebraically or directly from the definition.
- (a) $A \oplus (B \oplus C) = (A \oplus B) \oplus C$ (associative law for \oplus).
 - (b) $A \oplus \emptyset = A$.
 - (c) $A \oplus A^c = U$.
 - (d) $A \oplus A = \emptyset$.
 - (e) If $A \oplus C = B \oplus C$ then $A = B$.
- 1.14.** Let A , B , and C be subsets of U . Prove or disprove using Venn diagrams.
- (a) $A - B$ and $B - C$ are disjoint.
 - (b) $A - B$ and $C - B$ are disjoint.
 - (c) $A - (B \cup C)$ and $B - (A \cup C)$ are disjoint.
 - (d) $A - (B \cap C)$ and $B - (A \cap C)$ are disjoint.
- 1.15.** Which of the following are partitions of $\{1, 2, \dots, 8\}$? Explain your answers.
- (a) $\{\{1, 3, 5\}, \{1, 2, 6\}, \{4, 7, 8\}\}$
 - (b) $\{\{1, 3, 5\}, \{2, 6, 7\}, \{4, 8\}\}$
 - (c) $\{\{1, 3, 5\}, \{2, 6\}, \{2, 6\}, \{4, 7, 8\}\}$
 - (d) $\{\{1, 5\}, \{2, 6\}, \{4, 8\}\}$
- 1.16.** How many refinements are there of the partition $\{\{1, 3, 5\}, \{2, 6\}, \{4, 7, 8, 9\}\}$? Explain.
- 1.17.** Suppose S and T are sets with $S \cap T = \emptyset$. Suppose σ is a partition of S and τ is a partition of T .

- (a) Prove that $\sigma \cup \tau$ is a partition of $S \cup T$.
- (b) If σ has n_σ refinements and τ has n_τ refinements, how many refinements does $\sigma \cup \tau$ have? Explain.
- 1.18.** Use the characteristic function format to list the power set of the following sets. That is, describe each element of the power set as a vector of zeroes and ones.
- (a) $\{1, 2, 3\}$
- (b) $X \times Y$ where $X = \{a, b\}$ and $Y = \{x, y\}$.
- 1.19.** Find the following power sets:
- (a) $\mathcal{P}(\emptyset)$
- (b) $\mathcal{P}(\mathcal{P}(\emptyset))$
- (c) $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$
- 1.20.** Compare the following pairs of sets. Can they be equal? Is one a subset of the other? Can they have the same size (number of elements)?
- (a) $\mathcal{P}(A \cup B)$ and $\mathcal{P}(A) \cup \mathcal{P}(B)$
- (b) $\mathcal{P}(A \cap B)$ and $\mathcal{P}(A) \cap \mathcal{P}(B)$
- (c) $\mathcal{P}(A \times B)$ and $\mathcal{P}(A) \times \mathcal{P}(B)$
- 1.21.** Let $S = \{1, 2, \dots, n\}$. Let S_1 be the set of all subsets of S that contain 1. Let T_1 denote the set of all subsets of S that don't contain 1. Prove $|T_1| = |S_1| = 2^{(n-1)}$.

Section 2: Functions

Functions, such as linear functions, polynomial functions, trigonometric functions, exponential functions, and logarithmic functions are familiar to all students who have had mathematics in high school. For discrete mathematics, we need to understand functions at a basic set theoretic level. We begin with a familiar definition.

Definition 3 (Function) If A and B are sets, a *function* from A to B is a rule that tells us how to find a unique $b \in B$ for each $a \in A$. We write $f(a) = b$ and say that f maps a to b . We also say the value of f at a is b .

We write $f : A \rightarrow B$ to indicate that f is a function from A to B . We call the set A the *domain* of f and the set B the *range* or, equivalently, *codomain* of f .

To specify a function completely you must give its domain, range and rule.

If $X \subseteq A$, then $f(X) = \{f(x) \mid x \in X\}$. In particular $f(\emptyset) = \emptyset$ and $f(A)$ is called the *image* of f .

Sets and Functions

Some people define “range” to be the values that the function *actually* takes on. Most people call that the *image*.

In high school, you dealt with functions whose ranges were \mathbb{R} and whose domains were contained in \mathbb{R} ; for example, $f(x) = 1/(x^2 - 1)$ is a function from $\mathbb{R} - \{-1, 1\}$ to \mathbb{R} . If you have had some calculus, you also studied functions of functions! The derivative is a function whose domain is all differentiable functions and whose range is all functions. If we wanted to use functional notation we could write $D(f)$ to indicate the function that the derivative associates with f .

The set of all functions from A to B is written B^A . One reason for this notation, as we shall see below, is that $|B^A| = |B|^{|A|}$. Thus $f : A \rightarrow B$ and $f \in B^A$ say the same thing.

To avoid the cumbersome notation $\{1, 2, 3, \dots, n\}$, we will often use \underline{n} instead.

Example 10 (Functions as relations) There is a fundamental set-theoretic way of defining functions. Let A and B be sets. A *relation from A to B* is a subset of $A \times B$. For example, if $A = \underline{3} = \{1, 2, 3\}$ and $B = \underline{4}$, then $R = \{(1, 4), (1, 2), (3, 3), (2, 3)\}$ is a relation from A to B . To specify a relation, you must define three sets: A , B and R .

If the relation R satisfies the condition $\forall x \in A \exists! y \in B, (x, y) \in R$, then the relation R is called a *functional relation*. We used some shorthand notation here that is worth remembering:

\forall means “for all”
 \exists means “for some” or “there exists”
 $\exists!$ means “for exactly one”

If you think about Definition 3, you will realize that a “functional relation” is just one possible way of giving all of the information required to specify a function.

Given any relation $R \subseteq A \times B$, the inverse relation R^{-1} from B to A is $\{(y, x) : (x, y) \in R\}$. For $R = \{(1, 4), (1, 2), (3, 3), (2, 3)\}$, $A = \underline{3}$ and $B = \underline{4}$, the inverse relation is $R^{-1} = \{(4, 1), (2, 1), (3, 3), (3, 2)\}$. Note that neither R nor R^{-1} is a functional relation in this example. You should make sure that you understand why this statement is true. (Hint: R fails the “ $\exists!$ ” test and R^{-1} fails the “ \forall ” part of the definition of a functional relation.) Note also that if R and R^{-1} are functional then $|A| = |B|$. In algebra or calculus, when you draw a graph of a real-valued function $f : \mathbf{D} \rightarrow \mathbb{R}$ (such as $f(x) = x^3$, $f(x) = x/(1 - x)$, etc.), you are attempting a pictorial representation of the set $\{(x, f(x)) : x \in \mathbf{D} \subseteq \mathbb{R}\}$, which is the subset of $\mathbf{D} \times \mathbb{R}$. This subset is a “functional relation from \mathbf{D} to \mathbb{R} .”

In our notation, we would write $(a, b) \in R$ to indicate that the pair (a, b) is in the relation R from A to B . People also use the notation $a R b$ to indicate this. For example, the “less than” relation $\{(a, b) \mid a < b\}$ is written $a < b$. \blacksquare

In many cases in discrete mathematics, we are concerned with functions whose domain is finite. Special notation is used for specifying such functions.

Definition 4 (One-line notation) Let A be a finite ordered set with elements ordered $(a_1, a_2, \dots, a_{|A|})$. Let B be any set. A function $f : A \rightarrow B$ can be written in one-line notation as $f = (f(a_1), f(a_2), \dots, f(a_{|A|}))$. Thus the values of the function are written as

Section 2: Functions

list, which is also called a vector or a string. In other words the function f assigns to a_k the k^{th} element of the list $(f(a_1), f(a_2), \dots, f(a_{|A|}))$ for each value of k from 1 to $|A|$.

It follows from the definition that we can think of function as an element of $B^{|A|} = B \times B \times \dots \times B$, where there are $|A|$ copies of B . This is another reason for the notation B^A for all functions from A to B . Do you see why we don't use $B^{|A|}$ instead? No, it's not because B^A is easier to write. It's because $B^{|A|}$ does not specify the domain A . Instead, only its size $|A|$ is given.

Example 11 (Using the notation) To get a feeling for the notation used to specify a function, it may be helpful to imagine that you have an envelope or box that contains a function. In other words, this envelope contains all the information needed to completely describe the function. Think about what you're going to see when you open the envelope.

You might see

$$P = \{a, b, c\}, \quad g : P \rightarrow \underline{4}, \quad g(a) = 3, \quad g(b) = 1 \quad \text{and} \quad g(c) = 4.$$

This tells you that the name of the function is g , the domain of g is P , which is $\{a, b, c\}$, and the range of g is $\underline{4} = \{1, 2, 3, 4\}$. It also tells you the values in $\underline{4}$ that g assigns to each of the values in its domain. Someone else may have put

$$g \in \underline{4}^{\{a, b, c\}}, \quad \text{ordering: } a, b, c, \quad g = (3, 1, 4).$$

in the envelope instead. This describes the same function. It doesn't give a name for the domain, but that's okay since all we need to know is what's in the domain. On the other hand, it gives an order on the domain so that the function can be given in one-line form. Since the domain is ordered a, b, c and since $g = (3, 1, 4)$, by the definition of one-line notation $g(a) = 3$, $g(b) = 1$ and $g(c) = 4$. Can you describe other possible envelopes for the same function?

What if the envelope contained only $g = (3, 1, 4)$? If you think you have been given the one-line notation for g , you are mistaken. You *must* know the *ordered* domain of g before you can interpret $g = (3, 1, 4)$. Here we don't even know the domain as a set (or the range). The domain might be $\{a, b, c\}$, or $\{<, >, ?\}$, or any other 3-set.

What if the envelope contained

$$\text{the domain of } g \text{ is } \{a, b, c\}, \quad \text{ordering: } a, b, c, \quad g = (3, 1, 4)?$$

We haven't specified the range of g , but is it necessary since we know the values of the function? Our definition included the requirement that the range be specified, so this is not a complete definition. Some definitions of a function do not require that the range be specified. For such definitions, this would be a complete specification of the function g . \square

Sets and Functions

Example 12 (Counting functions) Think about specifying $f : A \rightarrow B$ in one-line notation: $(f(a_1), f(a_2), \dots, f(a_{|A|}))$. There are $|B|$ ways to choose $f(a_1)$, $|B|$ ways to choose $f(a_2)$, etc., and finally $|B|$ ways to choose $f(a_{|A|})$. This means that the cardinality of the set of all functions $f : A \rightarrow B$ is $|B|^{|A|}$. In other words, $|B^A| = |B|^{|A|}$.

We can represent a subset S of A by a unique function $f : A \rightarrow \underline{2}$ where

$$f(x) = \begin{cases} 1, & \text{if } x \notin S, \\ 2, & \text{if } x \in S. \end{cases}$$

This proves that there are $2^{|S|}$ such subsets. We proved this result in Example 9. You should verify that this is essentially the same proof that was given there.

We can represent a list of k elements of a set S with repetition allowed by a unique function $f : \underline{k} \rightarrow S$. In this representation, the list corresponds to the function written in one-line notation. (Recall that the ordering on \underline{k} is the numerical ordering.) This proves that there are exactly $|S|^k$ such lists. \square

Definition 5 (Types of functions) Let $f : A \rightarrow B$ be a function. If for every $b \in B$ there is an $a \in A$ such that $f(a) = b$, then f is called a *surjection* (or an *onto function*). Another way to describe a surjection is to say that it takes on each value in its range at least once.

If $f(x) = f(y)$ implies $x = y$, then f is called an *injection* (or a *one-to-one function*). Another way to describe an injection is to say that it takes on each value in its range at most once.

If f is both an injection and a surjection, it is called a *bijection*. The bijections of A^A are called the *permutations* of A . The set of permutations on a set A is denoted in various ways. Two notations are $\text{PER}(A)$ and $S(A)$.

If $f : A \rightarrow B$ is a bijection, we may talk about the *inverse bijection* of f , written f^{-1} , which reverses what f does. Thus $f^{-1} : B \rightarrow A$ and $f^{-1}(b)$ is that unique $a \in A$ such that $f(a) = b$.

Note that $f(f^{-1}(b)) = b$ and $f^{-1}(f(a)) = a$. Do not confuse f^{-1} with $1/f$. For example, if $f : \mathbb{R} \rightarrow \mathbb{R}$ is given by $f(x) = x^3 + 1$, then $1/f(x) = 1/(x^3 + 1)$ and $f^{-1}(x) = (x - 1)^{1/3}$.

Example 13 (Surjections, injections and bijections as lists) Lists provide another fundamental way to think about the various types of functions we've just defined. We'll illustrate this with some examples.

Let $A = \underline{4}$, $B = \{a, b, c, d, e\}$ and $f = (d, c, d, a)$ describe the function f in one-line notation. Since the list (d, c, d, a) contains d twice, f is not an injection. The function $(b, d, c, e) \in B^A$ is an injection since there are no repeats in the list of values taken on by the function. The 4-lists without repeats that can be formed from B correspond to the injections from $\underline{4}$ to B . In general, the injections in $S^{\underline{k}}$ correspond to k -lists without repeats whose elements are taken from S .

With the same f as in the previous paragraph, note that the value b is not taken on by f . Thus f is not a surjection. (We could have said e is not taken on, instead.)

Now let $A = \underline{4}$, $B = \{x, y, z\}$ and $g = (x, y, x, z)$. Since every element of B appears at least once in the list of values taken on, g is a surjection.

Finally, let $A = B = \underline{4}$ and $h = (3, 1, 4, 2)$. The function is both an injection and a surjection. Hence, it is a bijection. Since the domain and range are the same and f is a bijection, it is a permutation of $A = \underline{4}$. The list $(3, 1, 4, 2)$ is a rearrangement (a permutation) of the ordered listing $(1, 2, 3, 4)$ of A . That's why we call h a permutation. The inverse of h is $(2, 4, 1, 3)$. \square

Example 14 (Encryption) Suppose we want to send data (a text message, a JPEG file, etc.) to someone and want to be sure no one else can read the data. Then we use *encryption*. We can describe encryption as a function $f : D \rightarrow R$ where D is the set of possible messages. Of course, there are a huge number of *possible* messages, so what do we do? We can break the message into pieces. For example, we could break an ordinary text message into pieces with one character (with the space as a character) per piece. Then apply a function f to each piece. Here's a simple example: If x is a letter, $f(x)$ is the next letter in the alphabet with $Z \rightarrow A$ and $f(\text{space}) = \text{space}$. Then we would encrypt "HELLO THERE" as "IFMMP UIFSF." This is too simple for encryption.

What can we do? Let S be the set of symbols that we are using (A to Z and space in the previous paragraph). We could choose a more complicated function $f : S \rightarrow S$ than our simple function. What properties should it have?

- It must have an inverse so that we can decrypt.
- The encryption and decryption must be quick on a computer.
- It must be hard for someone else to figure out f^{-1} .

Since $f : S \rightarrow S$ and it has an inverse, f must be a bijection (in fact, a permutation of S). How can we make f hard to figure out? That is a problem in the design of encryption systems. One key ingredient is to make S large. For example, in systems like PGP (Pretty Good Privacy) and DES (Data Encryption Standard) S consists of all n -long vectors of zeroes and ones, typically with $n = 64$. In this case $|S| = 2^{64} \approx 10^{19}$, which is quite large. \square

Example 15 (Hashing) *Hashing* is a procedure for mapping a large space into a smaller one. For example, a hash function h may have as its domain all sequences of zeroes and ones of all possible lengths. Its range might be all n -long sequences of zeroes and ones for some n . There are some publicly available hash functions h that seem to be good.

Why would we want such a function? Suppose we want to be sure no one changes a document that is stored in a computer. We could apply h to the document and then save $h(\text{document})$. By giving $h(\text{document})$ to people, they could later check to see if the document had been changed — if the function h is well chosen it would be hard to change the document without changing the value of h , even if you know how to compute h . Suppose you email a document to a friend, but you're concerned that someone may intercept the email and change the document. You can call up your friend and tell him $h(\text{document})$ so that he can check it.

Another use for a hash function is storing data. Suppose we have an n -long array in which we want to store information about students at the university. We want a hash

Sets and Functions

function that maps student ID numbers into $\{1, 2, \dots, n\}$. Then $h(\text{ID})$ tells us which array position to use. Of course two student ID numbers may hash to the same value (array position). There are methods for dealing with such conflicts. \square

Example 16 (Two-line notation) Since one-line notation is a simple, brief way to specify functions, it is used frequently. If the domain is not a set of numbers, the notation is poor because we must first pause and order the domain. There are other ways to write functions which overcome this problem. For example, we could write $f(a) = 4$, $f(b) = 3$, $f(c) = 4$ and $f(d) = 1$. This could be shortened up somewhat to $a \rightarrow 4$, $b \rightarrow 3$, $c \rightarrow 4$ and $d \rightarrow 1$. By turning each of these sideways, we can shorten it even more: $\begin{pmatrix} a & b & c & d \\ 4 & 3 & 4 & 1 \end{pmatrix}$. For obvious reasons, this is called *two-line notation*. Since x always appears directly over $f(x)$, there is no need to order the domain; in fact, we need not even specify the domain separately since it is given by the top line. If the function is a bijection, its inverse is obtained by interchanging the top and bottom lines.

The arrows we introduced in the last paragraph can be used to help visualize different properties of functions. Imagine that you've listed the elements of the domain A in one column and the elements of the range B in another column to the right of the domain. Draw an arrow from a to b if $f(a) = b$. Thus the heads of arrows are on elements of B and the tails are on elements of A . Since f is a function, no two arrows have the same tail. If f is an injection, no two arrows have the same head. If f is a surjection, every element of B is on the head of some arrow. You should be able to describe the situation when f is a bijection. \square

Example 17 (Compositions of functions) Suppose that f and g are two functions such that the values f takes on are contained in the domain of g . We can write this as $f : A \rightarrow B$ and $g : C \rightarrow D$ where $f(a) \in C$ for all $a \in A$. We define the *composition* of g and f , written $gf : A \rightarrow D$ by $(gf)(x) = g(f(x))$ for all $x \in A$. The notation $g \circ f$ is also used to denote composition. Suppose that f and g are given in two-line notation by

$$f = \begin{pmatrix} p & q & r & s \\ P & R & T & U \end{pmatrix} \quad g = \begin{pmatrix} P & Q & R & S & T & U & V \\ 1 & 3 & 5 & 2 & 4 & 6 & 7 \end{pmatrix}.$$

Then $gf = \begin{pmatrix} p & q & r & s \\ 1 & 5 & 4 & 6 \end{pmatrix}.$

Suppose $f : A \rightarrow B$, $g : B \rightarrow C$, and $h : C \rightarrow D$. We can form the compositions $g \circ f$ and $h \circ g$; however, we cannot form the composition $h \circ f$ unless C contains $f(x)$ for all $x \in A$. We can also form the compositions of all three functions, namely $h \circ (g \circ f)$ and $(h \circ g) \circ f$. These two compositions are equal — that's the “associative law” for composition of functions. How is it proved? Here's an algebraic proof that uses nothing more than the definition of \circ at each step: For all $x \in A$

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x).$$

Let A be a set. Suppose that $f, g \in \mathcal{S}(A)$; that is, f and g are permutations of a set A . Recall that a permutation is a bijection from a set to itself and so it makes sense

Section 2: Functions

to talk about f^{-1} and fg . We claim that fg and f^{-1} are also permutations of A . This is easy to see if you write the permutations in two-line form and note that the second line is a rearrangement of the first if and only if the function is a permutation.

Again suppose that $f \in \mathcal{S}(A)$. Instead of $f \circ f$ or ff we write f^2 . Note that $f^2(x)$ is not $(f(x))^2$. (In fact, if multiplication is not defined in A , $(f(x))^2$ has no meaning.) We could compose three copies of f . The result is written f^3 . In general, we can compose k copies of f to obtain f^k . A cautious reader may be concerned that $f \circ (f \circ f)$ may not be the same as $(f \circ f) \circ f$. By the associative law for \circ , they're equal. In fact, $f^{k+m} = f^k \circ f^m$ for all nonnegative integers k and m , where f^0 is defined by $f^0(x) = x$ for all x in the domain. This is true even if k or m or both are negative. \square

Example 18 (Composing permutations) Let's carry out some calculations for practice. Let f and g be the permutations

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 5 & 3 \end{pmatrix} \quad g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}.$$

To compute fg , we must calculate $fg(x)$ for all x . This can be done fairly easily from the two-line form: For example, $(fg)(1)$ can be found by noting that the image of 1 under g is 2 and the image of 2 under f is 1. Thus $(fg)(1) = 1$. You should be able to verify that

$$fg = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 5 & 3 & 2 \end{pmatrix} \quad gf = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} \neq fg.$$

Thus, $f \circ g = g \circ f$ (commutative law) is *not a law* for permutations.

It is easy to get the inverse, simply interchange the two lines. Thus

$$f^{-1} = \begin{pmatrix} 2 & 1 & 4 & 5 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} \quad \text{which is the same as } f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 5 & 3 & 4 \end{pmatrix},$$

since the order of the columns in two-line form does not matter.

Let's compute some powers:

$$f^2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 & 4 \end{pmatrix} \quad f^3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{pmatrix} \quad g^5 = f^6 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

We computed f^6 using $f^6 = f^3 \circ f^3$. That was a bit tedious. Now imagine if you wanted to compute f^{100} . Cycle notation is an easy way to do that. \square

Let f be a permutation of the set A and let $n = |A|$. If $x \in A$, we can look at the sequence $x, f(x), f(f(x)), \dots, f^k(x), \dots$, which is often written as $x \rightarrow f(x) \rightarrow f(f(x)) \rightarrow \dots \rightarrow f^k(x) \rightarrow \dots$. Using the fact that $f^0(x) = x$, we can write the sequence as $f^0(x) \rightarrow f^1(x) \rightarrow f^2(x) \rightarrow \dots$. Since the range of f has n elements, this sequence will contain a repeated element in the first $n + 1$ entries. Suppose that $f^s(x)$ is the first sequence entry that is ever repeated and that $f^p(x)$ is the first time that it is repeated.

Sets and Functions

We claim that $s = 0$. If $s > 0$, apply f^{-1} to both sides of this equality to obtain $f^{s-1}(x) = f^{p-1}(x)$, contradicting the fact that s was chosen as small as possible. Thus, in fact, $s = 0$.

It follows that the sequence cycles through a pattern of length p forever since $f^{p+1}(x) = f(f^p(x)) = f(x)$, $f^{p+2}(x) = f^2(f^p(x)) = f^2(x)$, and so on. We call $(x, f(x), \dots, f^{p-1}(x))$ the *cycle* containing x and call p the *length of the cycle*. If a cycle has length p , we call it a p -cycle. Cyclic shifts of a cycle are considered the same; for example, if $(1, 2, 6, 3)$ is the cycle containing 1 (as well as 2, 3 and 6), then $(2, 6, 3, 1)$, $(6, 3, 1, 2)$ and $(3, 1, 2, 6)$ are other ways of writing the cycle.

Suppose (x_1, x_2, \dots, x_p) is a cycle of f and that $y_1 \in A$ is not in that cycle. We can form the cycle containing y_1 : (y_1, y_2, \dots, y_q) . None of the y_k is in the cycle (x_1, \dots, x_p) . Why? If it were, we could continue in the cycle and eventually reach y_1 . Written out algebraically: If $y_k = x_j$ for some k and j , then $y_1 = f^{q-k-1}(y_k) = f^{q-k-1}(x_j)$ and the right side is in the cycle (x_1, \dots, x_p) . We have proved.

Theorem 4 (Cycle form of a permutation) *Let f be a permutation of the finite set A be a finite set. Every element of A belongs to a cycle of f . Two cycles are either the same or have no elements in common.*

Example 19 (Using cycle notation) Consider the permutation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 8 & 1 & 5 & 9 & 3 & 7 & 6 \end{pmatrix}.$$

Since $1 \rightarrow 2 \rightarrow 4 \rightarrow 1$, the cycle containing 1 is $(1, 2, 4)$. We could equally well write it $(2, 4, 1)$ or $(4, 1, 2)$; however, $(1, 4, 2)$ is different since it corresponds to $1 \rightarrow 4 \rightarrow 2 \rightarrow 1$. The usual convention is to list the cycle starting with its smallest element. The cycles of f are $(1, 2, 4)$, $(3, 8, 7)$, (5) and $(6, 9)$. We write f in cycle form as

$$f = (1, 2, 4) (3, 8, 7) (5) (6, 9).$$

The order in which the cycles are written doesn't matter, so we have

$$f = (5) (6, 9) (1, 2, 4) (3, 8, 7) \quad \text{and} \quad f = (4, 1, 2) (5) (6, 9) (7, 3, 8),$$

and lots of other equivalent forms. It is common practice to omit the cycles of length one and write $f = (1, 2, 4)(3, 8, 7)(6, 9)$. The inverse of f is obtained by reading the cycles backwards because $f^{-1}(x)$ is the lefthand neighbor of x in a cycle. Thus

$$f^{-1} = (4, 2, 1)(7, 8, 3)(9, 6) = (1, 4, 2)(3, 7, 8)(6, 9).$$

To compute $f(x)$, we simply take one step to the right from x in its cycle. We just saw that $f^{-1}(x)$ is computed by taking one step to the left. You may be able to guess at this point that $f^k(x)$ is computed by taking k steps to the right, with the rule that a negative step to the right is the same as a positive step to the left. This makes it easy to compute powers.

Section 2: Functions

When $f = (1, 2, 4)(3, 8, 7)(5)(6, 9)$, what is f^{100} ? Imagine starting at 1. After 3 steps to the right we're back at 1. Do this 33 times so that after $3 \times 33 = 99$ steps to the right we're back at 1. One more step takes us to 100 steps and so $f^{100}(1) = 2$. You should be able to figure out the rest: $f^{100} = (1, 2, 4)(3, 8, 7)(5)(6)(9)$. \square

We next take a close look at the notions of *image* and *coimage* of a function. Again, let $f : A \rightarrow B$ be a function. The *image* of f is the set of values f actually takes on: $\text{Image}(f) = \{f(a) \mid a \in A\}$. The definition of a surjection can be rewritten $\text{Image}(f) = B$.

For each $b \in B$, the *inverse image* of b , written $f^{-1}(b)$ is the set of those elements in A whose image is b ; i.e.,

$$f^{-1}(b) = \{a \mid a \in A \text{ and } f(a) = b\}.$$

This extends our earlier definition of f^{-1} from bijections to all functions; however, such an f^{-1} can't be thought of as a function from B to A unless f is a bijection because it will not give a unique $a \in A$ for each $b \in B$. (There is a slight abuse of notation here: If $f : A \rightarrow B$ is a bijection, our new notation is $f^{-1}(b) = \{a\}$ and our old notation is $f^{-1}(b) = a$.)

Definition 6 (Coimage) Let $f : A \rightarrow B$ be a function. The collection of nonempty inverse images of elements of B is called the *coimage* of f . In set-theoretic terms,

$$\text{Coimage}(f) = \{f^{-1}(b) \mid b \in B, f^{-1}(b) \neq \emptyset\} = \{f^{-1}(b) \mid b \in \text{Image}(f)\}.$$

To describe the structure of coimages, we need to recall that a partition of a set is an unordered collection of nonempty subsets of B such that each element of B appears in exactly one subset. Each subset is called a block of the partition.

Theorem 5 (Structure of coimage) Suppose $f : A \rightarrow B$. The coimage of f is the partition of A whose blocks are the maximal subsets of A on which f is constant.

Wait — before we give a proof we need to understand what we just said. Let's look at an example. If $f \in \{a, b, c\}^5$ is given in one-line form as (a, c, a, a, c) , then

$$\text{Coimage}(f) = \{f^{-1}(a), f^{-1}(c)\} = \{\{1, 3, 4\}, \{2, 5\}\},$$

f is a on $\{1, 3, 4\}$ and is c on $\{2, 5\}$. Now let's prove the theorem.

Proof: If $x \in A$, let $y = f(x)$. Then $x \in f^{-1}(y)$ and so the union of the nonempty inverse images contains A . Clearly it does not contain anything which is not in A . If $y_1 \neq y_2$, then we cannot have $x \in f^{-1}(y_1)$ and $x \in f^{-1}(y_2)$ because this would imply $f(x) = y_1$ and $f(x) = y_2$, a contradiction of the definition of a function. Thus $\text{Coimage}(f)$ is a partition of A . Clearly x_1 and x_2 belong to the same block if and only if $f(x_1) = f(x_2)$. Hence a block is a maximal set on which f is constant. \square

Since $\text{Coimage}(f)$ is a partition of the domain A , we need to review the basic combinatorial properties of partitions.

Sets and Functions

Example 20 (Set partitions) The 15 partitions of $\{1, 2, 3, 4\}$, classified by number of blocks, are

1 block: $\{1, 2, 3, 4\}$
 2 blocks: $\{\{1, 2, 3\}, \{4\}\}$ $\{\{1, 2, 4\}, \{3\}\}$ $\{\{1, 2\}, \{3, 4\}\}$ $\{\{1, 3, 4\}, \{2\}\}$
 $\{\{1, 3\}, \{2, 4\}\}$ $\{\{1, 4\}, \{2, 3\}\}$ $\{\{1\}, \{2, 3, 4\}\}$
 3 blocks: $\{\{1, 2\}, \{3\}, \{4\}\}$ $\{\{1, 3\}, \{2\}, \{4\}\}$ $\{\{1, 4\}, \{2\}, \{3\}\}$ $\{\{1\}, \{2, 3\}, \{4\}\}$
 $\{\{1\}, \{2, 4\}, \{3\}\}$ $\{\{1\}, \{2\}, \{3, 4\}\}$
 4 blocks: $\{\{1\}, \{2\}, \{3\}, \{4\}\}$

Let $S(n, k)$ be the number of partitions of an n -set having exactly k blocks. These are called *Stirling numbers of the second kind*. Do not confuse $S(n, k)$ with $C(n, k) = \binom{n}{k}$. In both cases we have an n -set. For $C(n, k)$ we want to *choose a subset* containing k elements and for $S(n, k)$ we want to *partition the set* into k blocks.

What is the value of $S(n, k)$? Let's try to get a recursion. How can we build partitions of $\{1, 2, \dots, n\}$ with k blocks out of smaller cases? If we take partitions of $\{1, 2, \dots, n-1\}$ with $k-1$ blocks, we can simply add the block $\{n\}$. If we take partitions of $\{1, 2, \dots, n-1\}$ with k blocks, we can add the element n to one of the k blocks. You should convince yourself that all k block partitions of $\{1, 2, \dots, n\}$ arise in exactly one way when we do this. This gives us a recursion for $S(n, k)$. Putting n in a block by itself contributes $S(n-1, k-1)$. Putting n in a block with other elements contributes $S(n-1, k) \times k$. Thus,

$$S(n, k) = S(n-1, k-1) + k S(n-1, k).$$

Below is the tabular form for $S(n, k)$ analogous to the similar tabular form for $C(n, k)$.

| $n \backslash k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|------------------|---|----|----|----|----|----|---|
| 1 | 1 | | | | | | |
| 2 | 1 | 1 | | | | | |
| 3 | 1 | 3 | 1 | | | | |
| 4 | 1 | 7 | 6 | 1 | | | |
| 5 | 1 | 15 | 25 | 10 | 1 | | |
| 6 | 1 | 31 | 90 | 65 | 15 | 1 | |
| 7 | 1 | -- | -- | -- | -- | -- | 1 |

Notice that the starting conditions for this table are that $S(n, 1) = 1$ for all $n \geq 1$ and $S(n, n) = 1$ for all $n \geq 1$. The values for $n = 7$ are omitted from the table. You should fill them in to test your understanding of this computational process. For each n , the total number of partitions of a set of size n is equal to the sum $S(n, 1) + S(n, 2) + \dots + S(n, n)$. These numbers, gotten by summing the entries in the rows of the above table, are the Bell numbers, $B(n)$, that we discussed in Section 1. \square

Example 21 (Counting functions by image size) Suppose A and B are sets. Let $|A| = m$ and $|B| = n$. Suppose $k \leq m$ and $k \leq n$. A basic question about functions $f : A \rightarrow B$ is the following:

Let $S = \{f \mid f : A \rightarrow B, |\text{Image}(f)| = k\}$. Find $|S|$.

In other words, there are exactly k blocks in the coimage of f . This question clearly involves the Stirling numbers. In fact, the answer is $|S| = \binom{n}{k} S(m, k) k!$. The idea is to choose the image of the function in $\binom{n}{k}$ ways, then choose the coimage of the function in $S(m, k)$ ways and then put them together in $k!$ ways. You will get a chance to fill in the details in the last two exercises below. Here is an example. Suppose we take $|A| = 4$, $|B| = 5$, and $k = 3$. We get $|S| = \binom{5}{3} S(4, 3) 3! = 10 \times 6 \times 6 = 360$.

Let's look at some special cases.

- If $k = |B|$, then we are counting surjections. Why? We are given $|\text{Image}(f)| = |B|$. Thus every element in B must be in $\text{Image}(f)$.
- If $k = |A|$, then we are counting injections. Why? Suppose f is not an injection, say $f(a) = f(b)$ for some $a \neq b$. Then f can take on at most $|A| - 1$ different values. But $k = |A|$ says that $|\text{Image}(f)| = |A|$, a contradiction. \square

Exercises for Section 2

- 2.1.** Let R be the relation on $\mathbb{R} \times \mathbb{R}$, the Cartesian plane, defined by xRy if $y = x^2$. Sketch a picture that represents the set R in $\mathbb{R} \times \mathbb{R}$.
- 2.2.** Let R be a relation from the power set $\mathcal{P}(X)$ to itself, where $X = \{1, 2, 3, 4\}$, defined by ARB if $A \cap B \neq \emptyset$.
- (a) Is ARA for all $A \in \mathcal{P}(X)$?
 - (b) For any $A, B \in \mathcal{P}(X)$, if ARB is BRA ?
 - (c) For any $A, B \in \mathcal{P}(X)$, if ARB and BRC , is ARC ?
- 2.3.** In each case, draw the “directed graph diagram” of the given relation (label points in your diagram with the elements of S , put an arrow from x to y if and only if (x, y) belongs to the relation).
- (a) $S = \{(a, b), (a, c), (b, c), (d, d)\}$ on $X \times X$, $X = \{a, b, c, d\}$.
 - (b) Let $X = \{2, 3, 4, 5, 6, 7, 8\}$ and define xRy if $x = y \pmod{3}$; that is, if $|x - y|$ is divisible by 3.
- 2.4.** Find all relations on $\{a, b\} \times \{x, y\}$, that are not functional.
- 2.5.** Let S be the divides relation on $\{3, 4, 5\} \times \{4, 5, 6\}$; that is, xSy if y/x is an integer. List the elements of S and S^{-1} .
- 2.6.** Let A be a set with m elements and B be a set with n elements.

Sets and Functions

- (a) How many relations are there on $A \times B$?
- (b) How many functions are there from A to B ?

2.7. Define a binary relation D from $\underline{10} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ to $\underline{10}$ as follows: For all x, y , in $\underline{10}$, $x D y$ if $x < y$ and x divides y . How many edges are there in the directed graph of this relation. Explain.

2.8. This exercise lets you check your understanding of the definitions. In each case below, some information about a function is given to you. Answer the following questions and give reasons for your answers:

- Have you been given enough information to specify the function?
- Can you tell whether or not the function is an injection? a surjection? a bijection?
- If possible, give the function in two-line form.

- (a) $f \in \underline{3}^{\{>, <, +, ?\}}$, $f = (3, 1, 2, 3)$.
- (b) $f \in \{>, <, +, ?\}^3$, $f = (?, <, +)$.
- (c) $f \in \underline{4}^{\underline{3}}$, $2 \rightarrow 3$, $1 \rightarrow 4$, $3 \rightarrow 2$.

2.9. This exercise lets you check your understanding of the definitions. In each case below, some information about a function is given to you. Answer the following questions and give reasons for your answers:

- Have you been given enough information to specify the function?
- Can you tell whether or not the function is an injection? a surjection? a bijection? If so, what is it?

- (a) $f \in \underline{4}^{\underline{5}}$, $\text{Coimage}(f) = \{\{1, 3, 5\}, \{2, 4\}\}$.
- (b) $f \in \underline{5}^{\underline{5}}$, $\text{Coimage}(f) = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$.
- (c) $f \in \underline{4}^{\underline{5}}$, $f^{-1}(2) = \{1, 3, 5\}$, $f^{-1}(4) = \{2, 4\}$.
- (d) $f \in \underline{4}^{\underline{5}}$, $|\text{Image}(f)| = 4$.
- (e) $f \in \underline{4}^{\underline{5}}$, $|\text{Image}(f)| = 5$.
- (f) $f \in \underline{4}^{\underline{5}}$, $|\text{Coimage}(f)| = 5$.

2.10. For each of the following definitions, state whether the definition is correct or not correct. If not correct, explain why.

- (a) Definition: $f : A \rightarrow B$ is one-to-one if $\forall s, t \in A$, $f(s) = f(t)$ implies $s = t$.
- (b) Definition: $f : A \rightarrow B$ is one-to-one if $\forall s, t \in A$, $s \neq t$ implies $f(s) \neq f(t)$.
- (c) Definition: $f : A \rightarrow B$ is one-to-one if $\forall s \in A$, $\exists ! t \in B$ such that $f(s) = t$.

Section 2: Functions

- (d) Definition: $f : A \rightarrow B$ is one-to-one if $\forall t \in B, \exists$ at most one $s \in A$ such that $f(s) = t$.

2.11. Define $g : \mathbb{Z} \rightarrow \mathbb{Z}$ by $g(n) = 3n - 1$, where \mathbb{Z} is the set of integers.

- (a) Is g one-to-one?
- (b) Is g onto?
- (c) Suppose that $g : \mathbb{R} \rightarrow \mathbb{R}$ and $g(x) = 3x - 1$ for all real numbers x . Is g onto?

2.12. In each case prove or disprove the statement “ f is one-to-one.”

- (a) $f(x) = \frac{x}{x^2+1}$ where $f : \mathbb{R} \rightarrow \mathbb{R}$
- (b) $f(x) = \frac{2x+1}{x}$ where $f : (\mathbb{R} - \{0\}) \rightarrow \mathbb{R}$
- (c) $f(x) = \frac{x-1}{x+1}$ where $f : (\mathbb{R} - \{-1\}) \rightarrow \mathbb{R}$

2.13. This exercise lets you check your understanding of cycle form. A permutation is given in one-line, two-line or cycle form. Convert it to the other two forms. Give its inverse in all three forms.

- (a) (1,5,7,8) (2,3) (4) (6).
- (b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 3 & 7 & 2 & 6 & 4 & 5 & 1 \end{pmatrix}$.
- (c) (5,4,3,2,1), which is in one-line form.
- (d) (5,4,3,2,1), which is in cycle form. (Assume the domain is $\underline{5}$.)

2.14. Let $S = \{f \mid f : A \rightarrow B, f \text{ one-to-one}\}$. In each case, find $|S|$.

- (a) $|A| = 3$ and $|B| = 3$
- (b) $|A| = 3$ and $|B| = 5$
- (c) $|A| = m$ and $|B| = n$

2.15. Let $f : X \rightarrow Y$, $g : Y \rightarrow Z$ and $g \circ f : X \rightarrow Z$.

- (a) If $g \circ f$ is onto, must f and g be onto?
- (b) If $g \circ f$ is one-to-one, must f and g be one-to-one?

2.16. In each case, $f : X \rightarrow Y$ is an arbitrary function. Prove or disprove:

- (a) For all subsets A and B of X , $f(A \cup B) = f(A) \cup f(B)$.
- (b) For all subsets A and B of X , $f(A \cap B) = f(A) \cap f(B)$.
- (c) For all subsets A and B of X , $f(A - B) = f(A) - f(B)$.

Sets and Functions

(d) For all subsets C and D of Y , $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$.

2.17. Let $f : X \rightarrow Y$, $g : Y \rightarrow Z$ and $g \circ f : X \rightarrow Z$. Prove or disprove:

(a) For all subsets $A \subseteq X$, $f^{-1}(f(A)) = A$.

(b) For all subsets $B \subseteq Y$, $f(f^{-1}(B)) = B$.

(c) For all subsets $E \subseteq Z$, $(g \circ f)^{-1}(E) = f^{-1}(g^{-1}(E))$.

2.18. Let $S = \{f \mid f : A \rightarrow B, f \text{ onto}\}$. In each case, find $|S|$. Do (a)–(d) without using the general formula in the text.

(a) $|A| = 3$ and $|B| = 2$

(b) $|A| = 3$ and $|B| = 5$

(c) $|A| = 4$ and $|B| = 2$

(d) $|A| = m$ and $|B| = n$

(e) Explain how to use the general formula in the text to solve (d).

2.19. Let $S = \{f \mid f : A \rightarrow B, |\text{Image}(f)| = k\}$. Suppose that $|A| = m \geq k$ and $|B| = n \geq k$. Prove the formula for $|S|$ given in the text.

Multiple Choice Questions for Review

In each case there is one correct answer (given at the end of the problem set). Try to work the problem first without looking at the answer. Understand both why the correct answer is correct and why the other answers are wrong.

1. Which of the following statements is **FALSE**?
 - (a) $2 \in A \cup B$ implies that if $2 \notin A$ then $2 \in B$.
 - (b) $\{2, 3\} \subseteq A$ implies that $2 \in A$ and $3 \in A$.
 - (c) $A \cap B \supseteq \{2, 3\}$ implies that $\{2, 3\} \subseteq A$ and $\{2, 3\} \subseteq B$.
 - (d) $A - B \supseteq \{3\}$ and $\{2\} \subseteq B$ implies that $\{2, 3\} \subseteq A \cup B$.
 - (e) $\{2\} \in A$ and $\{3\} \in A$ implies that $\{2, 3\} \subseteq A$.
2. Let $A = \{0, 1\} \times \{0, 1\}$ and $B = \{a, b, c\}$. Suppose A is listed in lexicographic order based on $0 < 1$ and B is in alphabetic order. If $A \times B \times A$ is listed in lexicographic order, then the next element after $((1, 0), c, (1, 1))$ is
 - (a) $((1, 0), a, (0, 0))$
 - (b) $((1, 1), c, (0, 0))$
 - (c) $((1, 1), a, (0, 0))$
 - (d) $((1, 1), a, (1, 1))$
 - (e) $((1, 1), b, (1, 1))$
3. Which of the following statements is **TRUE**?
 - (a) For all sets A , B , and C , $A - (B - C) = (A - B) - C$.
 - (b) For all sets A , B , and C , $(A - B) \cap (C - B) = (A \cap C) - B$.
 - (c) For all sets A , B , and C , $(A - B) \cap (C - B) = A - (B \cup C)$.
 - (d) For all sets A , B , and C , if $A \cap C = B \cap C$ then $A = B$.
 - (e) For all sets A , B , and C , if $A \cup C = B \cup C$ then $A = B$.
4. Which of the following statements is **FALSE**?
 - (a) $C - (B \cup A) = (C - B) - A$
 - (b) $A - (C \cup B) = (A - B) - C$
 - (c) $B - (A \cup C) = (B - C) - A$
 - (d) $A - (B \cup C) = (B - C) - A$
 - (e) $A - (B \cup C) = (A - C) - B$
5. Consider the true theorem, "For all sets A and B , if $A \subseteq B$ then $A \cap B^c = \emptyset$." Which of the following statements is **NOT** equivalent to this statement:
 - (a) For all sets A^c and B , if $A \subseteq B$ then $A^c \cap B^c = \emptyset$.
 - (b) For all sets A and B , if $A^c \subseteq B$ then $A^c \cap B^c = \emptyset$.

Sets and Functions

- (c) For all sets A^c and B^c , if $A \subseteq B^c$ then $A \cap B = \emptyset$.
- (d) For all sets A^c and B^c , if $A^c \subseteq B^c$ then $A^c \cap B = \emptyset$.
- (e) For all sets A and B , if $A^c \supseteq B$ then $A \cap B = \emptyset$.
6. The power set $\mathcal{P}((A \times B) \cup (B \times A))$ has the same number of elements as the power set $\mathcal{P}((A \times B) \cup (A \times B))$ if and only if
- (a) $A = B$
- (b) $A = \emptyset$ or $B = \emptyset$
- (c) $B = \emptyset$ or $A = B$
- (d) $A = \emptyset$ or $B = \emptyset$ or $A = B$
- (e) $A = \emptyset$ or $B = \emptyset$ or $A \cap B = \emptyset$
7. Let $\sigma = 452631$ be a permutation on $\{1, 2, 3, 4, 5, 6\}$ in one-line notation (based on the usual order on integers). Which of the following is **NOT** a correct cycle notation for σ ?
- (a) $(614)(532)$
- (b) $(461)(352)$
- (c) $(253)(146)$
- (d) $(325)(614)$
- (e) $(614)(253)$
8. Let $f : X \rightarrow Y$. Consider the statement, “For all subsets C and D of Y , $f^{-1}(C \cap D^c) = f^{-1}(C) \cap [f^{-1}(D)]^c$. This statement is
- (a) True and equivalent to:
For all subsets C and D of Y , $f^{-1}(C - D) = f^{-1}(C) - f^{-1}(D)$.
- (b) False and equivalent to:
For all subsets C and D of Y , $f^{-1}(C - D) = f^{-1}(C) - f^{-1}(D)$.
- (c) True and equivalent to:
For all subsets C and D of Y , $f^{-1}(C - D) = f^{-1}(C) - [f^{-1}(D)]^c$.
- (d) False and equivalent to:
For all subsets C and D of Y , $f^{-1}(C - D) = f^{-1}(C) - [f^{-1}(D)]^c$.
- (e) True and equivalent to:
For all subsets C and D of Y , $f^{-1}(C - D) = [f^{-1}(C)]^c - f^{-1}(D)$.
9. Define $f(n) = \frac{n}{2} + \frac{1 - (-1)^n}{4}$ for all $n \in \mathbb{Z}$. Thus, $f : \mathbb{Z} \rightarrow \mathbb{Z}$, \mathbb{Z} the set of all integers. Which is correct?
- (a) f is not a function from $\mathbb{Z} \rightarrow \mathbb{Z}$ because $\frac{n}{2} \notin \mathbb{Z}$.
- (b) f is a function and is onto and one-to-one.
- (c) f is a function and is not onto but is one-to-one.
- (d) f is a function and is not onto and not one-to-one

- (e) f is a function and is onto but not one-to-one.
10. The number of partitions of $\{1, 2, 3, 4, 5\}$ into three blocks is $S(5, 3) = 25$. The total number of functions $f : \{1, 2, 3, 4, 5\} \rightarrow \{1, 2, 3, 4\}$ with $|\text{Image}(f)| = 3$ is
- (a) 4×6
 - (b) 4×25
 - (c) 25×6
 - (d) $4 \times 25 \times 6$
 - (e) $3 \times 25 \times 6$
11. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$. Let $h = g \circ f : X \rightarrow Z$. Suppose g is one-to-one and onto. Which of the following is **FALSE**?
- (a) If f is one-to-one then h is one-to-one and onto.
 - (b) If f is not onto then h is not onto.
 - (c) If f is not one-to-one then h is not one-to-one.
 - (d) If f is one-to-one then h is one-to-one.
 - (e) If f is onto then h is onto.
12. Which of the following statements is **FALSE**?
- (a) $\{2, 3, 4\} \subseteq A$ implies that $2 \in A$ and $\{3, 4\} \subseteq A$.
 - (b) $\{2, 3, 4\} \in A$ and $\{2, 3\} \in B$ implies that $\{4\} \subseteq A - B$.
 - (c) $A \cap B \supseteq \{2, 3, 4\}$ implies that $\{2, 3, 4\} \subseteq A$ and $\{2, 3, 4\} \subseteq B$.
 - (d) $A - B \supseteq \{3, 4\}$ and $\{1, 2\} \subseteq B$ implies that $\{1, 2, 3, 4\} \subseteq A \cup B$.
 - (e) $\{2, 3\} \subseteq A \cup B$ implies that if $\{2, 3\} \cap A = \emptyset$ then $\{2, 3\} \subseteq B$.
13. Let $A = \{0, 1\} \times \{0, 1\} \times \{0, 1\}$ and $B = \{a, b, c\} \times \{a, b, c\} \times \{a, b, c\}$. Suppose A is listed in lexicographic order based on $0 < 1$ and B is listed in lexicographic order based on $a < b < c$. If $A \times B \times A$ is listed in lexicographic order, then the next element after $((0, 1, 1), (c, c, c), (1, 1, 1))$ is
- (a) $((1, 0, 1), (a, a, b), (0, 0, 0))$
 - (b) $((1, 0, 0), (b, a, a), (0, 0, 0))$
 - (c) $((1, 0, 0), (a, a, a), (0, 0, 1))$
 - (d) $((1, 0, 0), (a, a, a), (1, 0, 0))$
 - (e) $((1, 0, 0), (a, a, a), (0, 0, 0))$
14. Consider the true theorem, "For all sets A , B , and C if $A \subseteq B \subseteq C$ then $C^c \subseteq B^c \subseteq A^c$." Which of the following statements is **NOT** equivalent to this statement:
- (a) For all sets A^c , B^c , and C^c , if $A^c \subseteq B^c \subseteq C^c$ then $C \subseteq B \subseteq A$.
 - (b) For all sets A^c , B , and C^c , if $A^c \subseteq B \subseteq C^c$ then $C \subseteq B^c \subseteq A$.
 - (c) For all sets A , B , and C^c , if $A^c \subseteq B \subseteq C$ then $C^c \subseteq B^c \subseteq A$.

Sets and Functions

- (d) For all sets A^c , B , and C^c , if $A^c \subseteq B^c \subseteq C$ then $C^c \subseteq B^c \subseteq A$.
- (e) For all sets A^c , B^c , and C^c , if $A^c \subseteq B^c \subseteq C$ then $C^c \subseteq B \subseteq A$.
- 15.** Let $\mathcal{P}(A)$ denote the power set of A . If $\mathcal{P}(A) \subseteq B$ then
- (a) $2^{|A|} \leq |B|$
- (b) $2^{|A|} \geq |B|$
- (c) $2|A| < |B|$
- (d) $|A| + 2 \leq |B|$
- (e) $2^{|A|} \geq 2^{|B|}$
- 16.** Let $f : \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \rightarrow \{a, b, c, d, e\}$. In one-line notation, $f = (e, a, b, b, a, c, c, a, c)$ (use number order on the domain). Which is correct?
- (a) $\text{Image}(f) = \{a, b, c, d, e\}$, $\text{Coimage}(f) = \{\{6, 7, 9\}, \{2, 5, 8\}, \{3, 4\}, \{1\}\}$
- (b) $\text{Image}(f) = \{a, b, c, e\}$, $\text{Coimage}(f) = \{\{6, 7, 9\}, \{2, 5, 8\}, \{3, 4\}\}$
- (c) $\text{Image}(f) = \{a, b, c, e\}$, $\text{Coimage}(f) = \{\{6, 7, 9\}, \{2, 5, 8\}, \{3, 4\}, \{1\}\}$
- (d) $\text{Image}(f) = \{a, b, c, e\}$, $\text{Coimage}(f) = \{\{6, 7, 9, 2, 5, 8\}, \{3, 4\}, \{1\}\}$
- (e) $\text{Image}(f) = \{a, b, c, d, e\}$, $\text{Coimage}(f) = \{\{1\}, \{3, 4\}, \{2, 5, 8\}, \{6, 7, 9\}\}$
- 17.** Let $\Sigma = \{x, y\}$ be an alphabet. The strings of length seven over Σ are listed in dictionary (lex) order. What is the first string after $xxxxyyx$ that is a palindrome (same read forwards and backwards)?
- (a) $xxxxxyx$ (b) $xxxyxxx$ (c) $xyxyxxx$ (d) $xyyyyxx$ (e) $xyxxxxyx$
- 18.** Let $\sigma = 681235947$ and $\tau = 627184593$ be permutations on $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ in one-line notation (based on the usual order on integers). Which of the following is a correct cycle notation for $\tau \circ \sigma$?
- (a) (124957368)
- (b) (142597368)
- (c) (142953768)
- (d) (142957368)
- (e) (142957386)

Answers: 1 (e), 2 (c), 3 (b), 4 (d), 5 (a), 6 (d), 7 (b), 8 (a), 9 (e), 10 (d), 11 (a), 12 (b), 13 (e), 14 (d), 15 (a), 16 (c), 17 (b), 18 (d).

We understand a “set” to be any collection M of certain distinct objects of our thought or intuition (called the “elements” of M) into a whole.
(Georg Cantor, 1895)

In mathematics you don’t understand things. You just get used to them.
(Attributed to John von Neumann)

In this chapter, we define sets, functions, and relations and discuss some of their general properties. This material can be referred back to as needed in the subsequent chapters.

1.2 Sets

A set is a collection of objects, called the elements or members of the set. The objects could be anything (planets, squirrels, characters in Shakespeare’s plays, or other sets) but for us they will be mathematical objects such as numbers, or sets of numbers. We write $x \in X$ if x is an element of the set X and $x \notin X$ if x is not an element of X .

If the definition of a “set” as a “collection” seems circular, that’s because it is. Conceiving of many objects as a single whole is a basic intuition that cannot be analyzed further, and the notions of “set” and “membership” are primitive ones. These notions can be made mathematically precise by introducing a system of axioms for sets and membership that agrees with our intuition and proving other set-theoretic properties from the axioms.

The most commonly used axioms for sets are the ZFC axioms, named somewhat inconsistently after two of their founders (Zermelo and Fraenkel) and one of their axioms (the Axiom of Choice). We won’t state these axioms here; instead, we use “naive” set theory, based on the intuitive properties of sets. Nevertheless, all the set-theory arguments we use can be rigorously formalized within the ZFC system.

Sets are determined entirely by their elements. Thus, the sets X , Y are equal, written $X = Y$, if

$$x \in X \text{ if and only if } x \in Y.$$

It is convenient to define the empty set, denoted by \emptyset , as the set with no elements. (Since sets are determined by their elements, there is only one set with no elements!) If $X \neq \emptyset$, meaning that X has at least one element, then we say that X is non-empty.

We can define a finite set by listing its elements (between curly brackets). For example,

$$X = \{2, 3, 5, 7, 11\}$$

is a set with five elements. The order in which the elements are listed or repetitions of the same element are irrelevant. Alternatively, we can define X as the set whose elements are the first five prime numbers. It doesn't matter how we specify the elements of X , only that they are the same.

Infinite sets can't be defined by explicitly listing all of their elements. Nevertheless, we will adopt a realist (or "platonist") approach towards arbitrary infinite sets and regard them as well-defined totalities. In constructive mathematics and computer science, one may be interested only in sets that can be defined by a rule or algorithm — for example, the set of all prime numbers — rather than by infinitely many arbitrary specifications, and there are some mathematicians who consider infinite sets to be meaningless without some way of constructing them. Similar issues arise with the notion of arbitrary subsets, functions, and relations.

1.1.1. Numbers. The infinite sets we use are derived from the natural and real numbers, about which we have a direct intuitive understanding.

Our understanding of the natural numbers $1, 2, 3, \dots$ derives from counting. We denote the set of natural numbers by

$$\mathbb{N} = \{1, 2, 3, \dots\}.$$

We define \mathbb{N} so that it starts at 1. In set theory and logic, the natural numbers are defined to start at zero, but we denote this set by $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Historically, the number 0 was later addition to the number system, primarily by Indian mathematicians in the 5th century AD. The ancient Greek mathematicians, such as Euclid, defined a number as a multiplicity and didn't consider 1 to be a number either.

Our understanding of the real numbers derives from durations of time and lengths in space. We think of the real line, or continuum, as being composed of an (uncountably) infinite number of points, each of which corresponds to a real number, and denote the set of real numbers by \mathbb{R} . There are philosophical questions, going back at least to Zeno's paradoxes, about whether the continuum can be represented as a set of points, and a number of mathematicians have disputed this assumption or introduced alternative models of the continuum. There are, however, no known inconsistencies in treating \mathbb{R} as a set of points, and since Cantor's work it has been the dominant point of view in mathematics because of its precision, power, and simplicity.

We denote the set of (positive, negative and zero) integers by

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\},$$

and the set of rational numbers (ratios of integers) by

$$\mathbb{Q} = \{p/q : p, q \in \mathbb{Z} \text{ and } q \neq 0\}.$$

The letter “Z” comes from “zahl” (German for “number”) and “Q” comes from “quotient.” These number systems are discussed further in Chapter 2.

Although we will not develop any complex analysis here, we occasionally make use of complex numbers. We denote the set of complex numbers by

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\},$$

where we add and multiply complex numbers in the natural way, with the additional identity that $i^2 = -1$, meaning that i is a square root of -1 . If $z = x + iy \in \mathbb{C}$, we call $x = \Re z$ the real part of z and $y = \Im z$ the imaginary part of z , and we call

$$|z| = \sqrt{x^2 + y^2}$$

the absolute value, or modulus, of z . Two complex numbers $z = x + iy$, $w = u + iv$ are equal if and only if $x = u$ and $y = v$.

1.1.2. Subsets. A set A is a subset of a set X , written $A \subset X$ or $X \supset A$, if every element of A belongs to X ; that is, if

$$x \in A \text{ implies that } x \in X.$$

We also say that A is included in X .¹ For example, if P is the set of prime numbers, then $P \subset \mathbb{N}$, and $\mathbb{N} \subset \mathbb{R}$. The empty set \emptyset and the whole set X are subsets of any set X . Note that $X = Y$ if and only if $X \subset Y$ and $Y \subset X$; we often prove the equality of two sets by showing that each one includes the other.

In our notation, $A \subset X$ does not imply that A is a proper subset of X (that is, a subset of X not equal to X itself), and we may have $A = X$. This notation for non-strict inclusion is not universal; some authors use $A \subset X$ to denote strict inclusion, in which $A \neq X$, and $A \subseteq X$ to denote non-strict inclusion, in which $A = X$ is allowed.

Definition 1.1. The power set $\mathcal{P}(X)$ of a set X is the set of all subsets of X .

Example 1.2. If $X = \{1, 2, 3\}$, then

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{2, 3\}, \{1, 3\}, \{1, 2\}, \{1, 2, 3\}\}.$$

The power set of a finite set with n elements has 2^n elements because, in defining a subset, we have two independent choices for each element (does it belong to the subset or not?). In Example 1.2, X has 3 elements and $\mathcal{P}(X)$ has $2^3 = 8$ elements.

The power set of an infinite set, such as \mathbb{N} , consists of all finite and infinite subsets and is infinite. We can define finite subsets of \mathbb{N} , or subsets with finite

¹By contrast, we say that an element $x \in X$ is *contained* in X , in which cases the singleton set $\{x\}$ is *included* in X . This terminological distinction is not universal, but it is almost always clear from the context whether one is referring to an element of a set or a subset of a set. In fact, before the development of the contemporary notation for set theory, Dedekind [3] used the same symbol (\subseteq) to denote both membership of elements and inclusion of subsets.

complements, by listing finitely many elements. Some infinite subsets, such as the set of primes or the set of squares, can be defined by giving a definite rule for membership. We imagine that a general subset $A \subset \mathbb{N}$ is “defined” by going through the elements of \mathbb{N} one by one and deciding for each $n \in \mathbb{N}$ whether $n \in A$ or $n \notin A$.

If X is a set and P is a property of elements of X , we denote the subset of X consisting of elements with the property P by $\{x \in X : P(x)\}$.

Example 1.3. The set

$$\{n \in \mathbb{N} : n = k^2 \text{ for some } k \in \mathbb{N}\}$$

is the set of perfect squares $\{1, 4, 9, 16, 25, \dots\}$. The set

$$\{x \in \mathbb{R} : 0 < x < 1\}$$

is the open interval $(0, 1)$.

1.1.3. Set operations. The intersection $A \cap B$ of two sets A, B is the set of all elements that belong to both A and B ; that is

$$x \in A \cap B \text{ if and only if } x \in A \text{ and } x \in B.$$

Two sets A, B are said to be disjoint if $A \cap B = \emptyset$; that is, if A and B have no elements in common.

The union $A \cup B$ is the set of all elements that belong to A or B ; that is

$$x \in A \cup B \text{ if and only if } x \in A \text{ or } x \in B.$$

Note that we always use ‘or’ in an inclusive sense, so that $x \in A \cup B$ if x is an element of A or B , or both A and B . (Thus, $A \cap B \subset A \cup B$.)

The set-difference of two sets B and A is the set of elements of B that do not belong to A ,

$$B \setminus A = \{x \in B : x \notin A\}.$$

If we consider sets that are subsets of a fixed set X that is understood from the context, then we write $A^c = X \setminus A$ to denote the complement of $A \subset X$ in X . Note that $(A^c)^c = A$.

Example 1.4. If

$$A = \{2, 3, 5, 7, 11\}, \quad B = \{1, 3, 5, 7, 9, 11\}$$

then

$$A \cap B = \{3, 5, 7, 11\}, \quad A \cup B = \{1, 2, 3, 5, 7, 9, 11\}.$$

Thus, $A \cap B$ consists of the natural numbers between 1 and 11 that are both prime and odd, while $A \cup B$ consists of the numbers that are either prime or odd (or both). The set differences of these sets are

$$B \setminus A = \{1, 9\}, \quad A \setminus B = \{2\}.$$

Thus, $B \setminus A$ is the set of odd numbers between 1 and 11 that are not prime, and $A \setminus B$ is the set of prime numbers that are not odd.

These set operations may be represented by Venn diagrams, which can be used to visualize their properties. In particular, if $A, B \subset X$, we have De Morgan's laws:

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c.$$

The definitions of union and intersection extend to larger collections of sets in a natural way.

Definition 1.5. Let \mathcal{C} be a collection of sets. Then the union of \mathcal{C} is

$$\bigcup \mathcal{C} = \{x : x \in X \text{ for some } X \in \mathcal{C}\},$$

and the intersection of \mathcal{C} is

$$\bigcap \mathcal{C} = \{x : x \in X \text{ for every } X \in \mathcal{C}\}.$$

If $\mathcal{C} = \{A, B\}$, then this definition reduces to our previous one for $A \cup B$ and $A \cap B$.

The Cartesian product $X \times Y$ of sets X, Y is the set of all ordered pairs (x, y) with $x \in X$ and $y \in Y$. If $X = Y$, we often write $X \times X = X^2$. Two ordered pairs $(x_1, y_1), (x_2, y_2)$ in $X \times Y$ are equal if and only if $x_1 = x_2$ and $y_1 = y_2$. Thus, $(x, y) \neq (y, x)$ unless $x = y$. This contrasts with sets where $\{x, y\} = \{y, x\}$.

Example 1.6. If $X = \{1, 2, 3\}$ and $Y = \{4, 5\}$ then

$$X \times Y = \{(1, 4), (1, 5), (2, 4), (2, 5), (3, 4), (3, 5)\}.$$

Example 1.7. The Cartesian product of \mathbb{R} with itself is the Cartesian plane \mathbb{R}^2 consisting of all points with coordinates (x, y) where $x, y \in \mathbb{R}$.

The Cartesian product of finitely many sets is defined analogously.

Definition 1.8. The Cartesian products of n sets X_1, X_2, \dots, X_n is the set of ordered n -tuples,

$$X_1 \times X_2 \times \cdots \times X_n = \{(x_1, x_2, \dots, x_n) : x_i \in X_i \text{ for } i = 1, 2, \dots, n\},$$

where $(x_1, x_2, \dots, x_n) = (y_1, y_2, \dots, y_n)$ if and only if $x_i = y_i$ for every $i = 1, 2, \dots, n$.

1.2. Functions

A function $f : X \rightarrow Y$ between sets X, Y assigns to each $x \in X$ a unique element $f(x) \in Y$. Functions are also called maps, mappings, or transformations. The set X on which f is defined is called the domain of f and the set Y in which it takes its values is called the codomain. We write $f : x \mapsto f(x)$ to indicate that f is the function that maps x to $f(x)$.

Example 1.9. The identity function $\text{id}_X : X \rightarrow X$ on a set X is the function $\text{id}_X : x \mapsto x$ that maps every element to itself.

Example 1.10. Let $A \subset X$. The characteristic (or indicator) function of A ,

$$\chi_A : X \rightarrow \{0, 1\},$$

is defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Specifying the function χ_A is equivalent to specifying the subset A .

Example 1.11. Let A, B be the sets in Example 1.4. We can define a function $f : A \rightarrow B$ by

$$f(2) = 7, \quad f(3) = 1, \quad f(5) = 11, \quad f(7) = 3, \quad f(11) = 9,$$

and a function $g : B \rightarrow A$ by

$$g(1) = 3, \quad g(3) = 7, \quad g(5) = 2, \quad g(7) = 2, \quad g(9) = 5, \quad g(11) = 11.$$

Example 1.12. The square function $f : \mathbb{N} \rightarrow \mathbb{N}$ is defined by

$$f(n) = n^2,$$

which we also write as $f : n \mapsto n^2$. The equation $g(n) = \sqrt{n}$, where \sqrt{n} is the positive square root, defines a function $g : \mathbb{N} \rightarrow \mathbb{R}$, but $h(n) = \pm\sqrt{n}$ does not define a function since it doesn't specify a unique value for $h(n)$. Sometimes we use a convenient oxymoron and refer to h as a multi-valued function.

One way to specify a function is to explicitly list its values, as in Example 1.11. Another way is to give a definite rule, as in Example 1.12. If X is infinite and f is not given by a definite rule, then neither of these methods can be used to specify the function. Nevertheless, we suppose that a general function $f : X \rightarrow Y$ may be “defined” by picking for each $x \in X$ a corresponding value $f(x) \in Y$.

If $f : X \rightarrow Y$ and $U \subset X$, then we denote the restriction of f to U by $f|_U : U \rightarrow Y$, where $f|_U(x) = f(x)$ for $x \in U$.

In defining a function $f : X \rightarrow Y$, it is crucial to specify the domain X of elements on which it is defined. There is more ambiguity about the choice of codomain, however, since we can extend the codomain to any set $Z \supset Y$ and define a function $g : X \rightarrow Z$ by $g(x) = f(x)$. Strictly speaking, even though f and g have exactly the same values, they are different functions since they have different codomains. Usually, however, we will ignore this distinction and regard f and g as being the same function.

The graph of a function $f : X \rightarrow Y$ is the subset G_f of $X \times Y$ defined by

$$G_f = \{(x, y) \in X \times Y : x \in X \text{ and } y = f(x)\}.$$

For example, if $f : \mathbb{R} \rightarrow \mathbb{R}$, then the graph of f is the usual set of points (x, y) with $y = f(x)$ in the Cartesian plane \mathbb{R}^2 . Since a function is defined at every point in its domain, there is some point $(x, y) \in G_f$ for every $x \in X$, and since the value of a function is uniquely defined, there is exactly one such point. In other words, for each $x \in X$ the “vertical line” $L_x = \{(x, y) \in X \times Y : y \in Y\}$ through x intersects the graph of a function $f : X \rightarrow Y$ in exactly one point: $L_x \cap G_f = (x, f(x))$.

Definition 1.13. The range, or image, of a function $f : X \rightarrow Y$ is the set of values

$$\text{ran } f = \{y \in Y : y = f(x) \text{ for some } x \in X\}.$$

A function is onto if its range is all of Y ; that is, if

$$\text{for every } y \in Y \text{ there exists } x \in X \text{ such that } y = f(x).$$

A function is one-to-one if it maps distinct elements of X to distinct elements of Y ; that is, if

$$x_1, x_2 \in X \text{ and } x_1 \neq x_2 \text{ implies that } f(x_1) \neq f(x_2).$$

An onto function is also called a surjection, a one-to-one function an injection, and a one-to-one, onto function a bijection.

Example 1.14. The function $f : A \rightarrow B$ defined in Example 1.11 is one-to-one but not onto, since $5 \notin \text{ran } f$, while the function $g : B \rightarrow A$ is onto but not one-to-one, since $g(5) = g(7)$.

1.3. Composition and inverses of functions

The successive application of mappings leads to the notion of the composition of functions.

Definition 1.15. The composition of functions $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ is the function $g \circ f : X \rightarrow Z$ defined by

$$(g \circ f)(x) = g(f(x)).$$

The order of application of the functions in a composition is crucial and is read from right to left. The composition $g \circ f$ can only be defined if the domain of g includes the range of f , and the existence of $g \circ f$ does not imply that $f \circ g$ even makes sense.

Example 1.16. Let X be the set of students in a class and $f : X \rightarrow \mathbb{N}$ the function that maps a student to her age. Let $g : \mathbb{N} \rightarrow \mathbb{N}$ be the function that adds up the digits in a number e.g., $g(1729) = 19$. If $x \in X$ is 23 years old, then $(g \circ f)(x) = 5$, but $(f \circ g)(x)$ makes no sense, since students in the class are not natural numbers.

Even if both $g \circ f$ and $f \circ g$ are defined, they are, in general, different functions.

Example 1.17. If $f : A \rightarrow B$ and $g : B \rightarrow A$ are the functions in Example 1.11, then $g \circ f : A \rightarrow A$ is given by

$$\begin{aligned} (g \circ f)(2) &= 2, & (g \circ f)(3) &= 3, & (g \circ f)(5) &= 11, \\ (g \circ f)(7) &= 7, & (g \circ f)(11) &= 5. \end{aligned}$$

and $f \circ g : B \rightarrow B$ is given by

$$\begin{aligned} (f \circ g)(1) &= 1, & (f \circ g)(3) &= 3, & (f \circ g)(5) &= 7, \\ (f \circ g)(7) &= 7, & (f \circ g)(9) &= 11, & (f \circ g)(11) &= 9. \end{aligned}$$

A one-to-one, onto function $f : X \rightarrow Y$ has an inverse $f^{-1} : Y \rightarrow X$ defined by

$$f^{-1}(y) = x \text{ if and only if } f(x) = y.$$

Equivalently, $f^{-1} \circ f = \text{id}_X$ and $f \circ f^{-1} = \text{id}_Y$. A value $f^{-1}(y)$ is defined for every $y \in Y$ since f is onto, and it is unique since f is one-to-one. If $f : X \rightarrow Y$ is one-to-one but not onto, then one can still define an inverse function $f^{-1} : \text{ran } f \rightarrow X$ whose domain is the range of f .

The use of the notation f^{-1} to denote the inverse function should not be confused with its use to denote the reciprocal function; it should be clear from the context which meaning is intended.

Example 1.18. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is the function $f(x) = x^3$, which is one-to-one and onto, then the inverse function $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ is given by

$$f^{-1}(x) = x^{1/3}.$$

On the other hand, the reciprocal function $g = 1/f$ is given by

$$g(x) = \frac{1}{x^3}, \quad g : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}.$$

The reciprocal function is not defined at $x = 0$ where $f(x) = 0$.

If $f : X \rightarrow Y$ and $A \subset X$, then we let

$$f(A) = \{y \in Y : y = f(x) \text{ for some } x \in A\}$$

denote the set of values of f on points in A . Similarly, if $B \subset Y$, we let

$$f^{-1}(B) = \{x \in X : f(x) \in B\}$$

denote the set of points in X whose values belong to B . Note that $f^{-1}(B)$ makes sense as a set even if the inverse function $f^{-1} : Y \rightarrow X$ does not exist.

Example 1.19. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^2$. If $A = (-2, 2)$, then $f(A) = [0, 4)$. If $B = (0, 4)$, then

$$f^{-1}(B) = (-2, 0) \cup (0, 2).$$

If $C = (-4, 0)$, then $f^{-1}(C) = \emptyset$.

Finally, we introduce operations on a set.

Definition 1.20. A binary operation on a set X is a function $f : X \times X \rightarrow X$.

We think of f as “combining” two elements of X to give another element of X . One can also consider higher-order operations, such as ternary operations $f : X \times X \times X \rightarrow X$, but will only use binary operations.

Example 1.21. Addition $a : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ and multiplication $m : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ are binary operations on \mathbb{N} where

$$a(x, y) = x + y, \quad m(x, y) = xy.$$

1.4. Indexed sets

We say that a set X is indexed by a set I , or X is an indexed set, if there is an onto function $f : I \rightarrow X$. We then write

$$X = \{x_i : i \in I\}$$

where $x_i = f(i)$. For example,

$$\{1, 4, 9, 16, \dots\} = \{n^2 : n \in \mathbb{N}\}.$$

The set X itself is the range of the indexing function f , and it doesn't depend on how we index it. If f isn't one-to-one, then some elements are repeated, but this doesn't affect the definition of the set X . For example,

$$\{-1, 1\} = \{(-1)^n : n \in \mathbb{N}\} = \{(-1)^{n+1} : n \in \mathbb{N}\}.$$

If $\mathcal{C} = \{X_i : i \in I\}$ is an indexed collection of sets X_i , then we denote the union and intersection of the sets in \mathcal{C} by

$$\bigcup_{i \in I} X_i = \{x : x \in X_i \text{ for some } i \in I\}, \quad \bigcap_{i \in I} X_i = \{x : x \in X_i \text{ for every } i \in I\},$$

or similar notation.

Example 1.22. For $n \in \mathbb{N}$, define the intervals

$$A_n = [1/n, 1 - 1/n] = \{x \in \mathbb{R} : 1/n \leq x \leq 1 - 1/n\},$$

$$B_n = (-1/n, 1/n) = \{x \in \mathbb{R} : -1/n < x < 1/n\}.$$

Then

$$\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n=1}^{\infty} A_n = (0, 1), \quad \bigcap_{n \in \mathbb{N}} B_n = \bigcap_{n=1}^{\infty} B_n = \{0\}.$$

The general statement of De Morgan's laws for a collection of sets is as follows.

Proposition 1.23 (De Morgan). If $\{X_i \subset X : i \in I\}$ is a collection of subsets of a set X , then

$$\left(\bigcup_{i \in I} X_i \right)^c = \bigcap_{i \in I} X_i^c, \quad \left(\bigcap_{i \in I} X_i \right)^c = \bigcup_{i \in I} X_i^c.$$

Proof. We have $x \notin \bigcup_{i \in I} X_i$ if and only if $x \notin X_i$ for every $i \in I$, which holds if and only if $x \in \bigcap_{i \in I} X_i^c$. Similarly, $x \notin \bigcap_{i \in I} X_i$ if and only if $x \notin X_i$ for some $i \in I$, which holds if and only if $x \in \bigcup_{i \in I} X_i^c$. \square

The following theorem summarizes how unions and intersections map under functions.

Theorem 1.24. Let $f : X \rightarrow Y$ be a function. If $\{Y_j \subset Y : j \in J\}$ is a collection of subsets of Y , then

$$f^{-1} \left(\bigcup_{j \in J} Y_j \right) = \bigcup_{j \in J} f^{-1}(Y_j), \quad f^{-1} \left(\bigcap_{j \in J} Y_j \right) = \bigcap_{j \in J} f^{-1}(Y_j);$$

and if $\{X_i \subset X : i \in I\}$ is a collection of subsets of X , then

$$f \left(\bigcup_{i \in I} X_i \right) = \bigcup_{i \in I} f(X_i), \quad f \left(\bigcap_{i \in I} X_i \right) \subset \bigcap_{i \in I} f(X_i).$$

Proof. We prove only the results for the inverse image of a union and the image of an intersection; the proof of the remaining two results is similar.

If $x \in f^{-1} \left(\bigcup_{j \in J} Y_j \right)$, then there exists $y \in \bigcup_{j \in J} Y_j$ such that $f(x) = y$. Then $y \in Y_j$ for some $j \in J$ and $x \in f^{-1}(Y_j)$, so $x \in \bigcup_{j \in J} f^{-1}(Y_j)$. It follows that

$$f^{-1} \left(\bigcup_{j \in J} Y_j \right) \subset \bigcup_{j \in J} f^{-1}(Y_j).$$

Conversely, if $x \in \bigcup_{j \in J} f^{-1}(Y_j)$, then $x \in f^{-1}(Y_j)$ for some $j \in J$, so $f(x) \in Y_j$ and $f(x) \in \bigcup_{j \in J} Y_j$, meaning that $x \in f^{-1}\left(\bigcup_{j \in J} Y_j\right)$. It follows that

$$\bigcup_{j \in J} f^{-1}(Y_j) \subset f^{-1}\left(\bigcup_{j \in J} Y_j\right),$$

which proves that the sets are equal.

If $y \in f\left(\bigcap_{i \in I} X_i\right)$, then there exists $x \in \bigcap_{i \in I} X_i$ such that $f(x) = y$. Then $x \in X_i$ and $y \in f(X_i)$ for every $i \in I$, meaning that $y \in \bigcap_{i \in I} f(X_i)$. It follows that

$$f\left(\bigcap_{i \in I} X_i\right) \subset \bigcap_{i \in I} f(X_i).$$

□

The only case in which we don't always have equality is for the image of an intersection, and we may get strict inclusion here if f is not one-to-one.

Example 1.25. Define $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^2$. Let $A = (-1, 0)$ and $B = (0, 1)$. Then $A \cap B = \emptyset$ and $f(A \cap B) = \emptyset$, but $f(A) = f(B) = (0, 1)$, so $f(A) \cap f(B) = (0, 1) \neq f(A \cap B)$.

Next, we generalize the Cartesian product of finitely many sets to the product of possibly infinitely many sets.

Definition 1.26. Let $\mathcal{C} = \{X_i : i \in I\}$ be an indexed collection of sets X_i . The Cartesian product of \mathcal{C} is the set of functions that assign to each index $i \in I$ an element $x_i \in X_i$. That is,

$$\prod_{i \in I} X_i = \left\{ f : I \rightarrow \bigcup_{i \in I} X_i : f(i) \in X_i \text{ for every } i \in I \right\}.$$

For example, if $I = \{1, 2, \dots, n\}$, then f defines an ordered n -tuple of elements (x_1, x_2, \dots, x_n) with $x_i = f(i) \in X_i$, so this definition is equivalent to our previous one.

If $X_i = X$ for every $i \in I$, then $\prod_{i \in I} X_i$ is simply the set of functions from I to X , and we also write it as

$$X^I = \{f : I \rightarrow X\}.$$

We can think of this set as the set of ordered I -tuples of elements of X .

Example 1.27. A sequence of real numbers $(x_1, x_2, x_3, \dots, x_n, \dots) \in \mathbb{R}^{\mathbb{N}}$ is a function $f : \mathbb{N} \rightarrow \mathbb{R}$. We study sequences and their convergence properties in Chapter 3.

Example 1.28. Let $\mathbf{2} = \{0, 1\}$ be a set with two elements. Then a subset $A \subset I$ can be identified with its characteristic function $\chi_A : I \rightarrow \mathbf{2}$ by: $i \in A$ if and only if $\chi_A(i) = 1$. Thus, $A \mapsto \chi_A$ is a one-to-one map from $\mathcal{P}(I)$ onto $\mathbf{2}^I$.

Before giving another example, we introduce some convenient notation.

Definition 1.29. Let

$$\Sigma = \{(s_1, s_2, s_3, \dots, s_k, \dots) : s_k = 0, 1\}$$

denote the set of all binary sequences; that is, sequences whose terms are either 0 or 1.

Example 1.30. Let $\mathbf{2} = \{0, 1\}$. Then $\Sigma = \mathbf{2}^{\mathbb{N}}$, where we identify a sequence $(s_1, s_2, \dots, s_k, \dots)$ with the function $f : \mathbb{N} \rightarrow \mathbf{2}$ such that $s_k = f(k)$. We can also identify Σ and $\mathbf{2}^{\mathbb{N}}$ with $\mathcal{P}(\mathbb{N})$ as in Example 1.28. For example, the sequence $(1, 0, 1, 0, 1, \dots)$ of alternating ones and zeros corresponds to the function $f : \mathbb{N} \rightarrow \mathbf{2}$ defined by

$$f(k) = \begin{cases} 1 & \text{if } k \text{ is odd,} \\ 0 & \text{if } k \text{ is even,} \end{cases}$$

and to the set $\{1, 3, 5, 7, \dots\} \subset \mathbb{N}$ of odd natural numbers.

1.5. Relations

A binary relation R on sets X and Y is a definite relation between elements of X and elements of Y . We write xRy if $x \in X$ and $y \in Y$ are related. One can also define relations on more than two sets, but we shall consider only binary relations and refer to them simply as relations. If $X = Y$, then we call R a relation on X .

Example 1.31. Suppose that S is a set of students enrolled in a university and B is a set of books in a library. We might define a relation R on S and B by:

$$s \in S \text{ has read } b \in B.$$

In that case, sRb if and only if s has read b . Another, probably inequivalent, relation is:

$$s \in S \text{ has checked } b \in B \text{ out of the library.}$$

When used informally, relations may be ambiguous (did s read b if she only read the first page?), but in mathematical usage we always require that relations are definite, meaning that one and only one of the statements “these elements are related” or “these elements are not related” is true.

The graph G_R of a relation R on X and Y is the subset of $X \times Y$ defined by

$$G_R = \{(x, y) \in X \times Y : xRy\}.$$

This graph contains all of the information about which elements are related. Conversely, any subset $G \subset X \times Y$ defines a relation R by: xRy if and only if $(x, y) \in G$. Thus, a relation on X and Y may be (and often is) defined as subset of $X \times Y$. As for sets, it doesn’t matter how a relation is defined, only what elements are related.

A function $f : X \rightarrow Y$ determines a relation F on X and Y by: xFy if and only if $y = f(x)$. Thus, functions are a special case of relations. The graph G_R of a general relation differs from the graph G_F of a function in two ways: there may be elements $x \in X$ such that $(x, y) \notin G_R$ for any $y \in Y$, and there may be $x \in X$ such that $(x, y) \in G_R$ for many $y \in Y$.

For example, in the case of the relation R in Example 1.31, there may be some students who haven’t read any books, and there may be other students who have

read lots of books, in which case we don't have a well-defined function from students to books.

Two important types of relations are orders and equivalence relations, and we define them next.

1.5.1. Orders. A primary example of an order is the standard order \leq on the natural (or real) numbers. This order is a linear or total order, meaning that two numbers are always comparable. Another example of an order is inclusion \subset on the power set of some set; one set is “smaller” than another set if it is included in it. This order is a partial order (provided the original set has at least two elements), meaning that two subsets need not be comparable.

Example 1.32. Let $X = \{1, 2\}$. The collection of subsets of X is

$$\mathcal{P}(X) = \{\emptyset, A, B, X\}, \quad A = \{1\}, \quad B = \{2\}.$$

We have $\emptyset \subset A \subset X$ and $\emptyset \subset B \subset X$, but $A \not\subset B$ and $B \not\subset A$, so A and B are not comparable under ordering by inclusion.

The general definition of an order is as follows.

Definition 1.33. An order \preceq on a set X is a binary relation on X such that for every $x, y, z \in X$:

- (a) $x \preceq x$ (reflexivity);
- (b) if $x \preceq y$ and $y \preceq x$ then $x = y$ (antisymmetry);
- (c) if $x \preceq y$ and $y \preceq z$ then $x \preceq z$ (transitivity).

An order is a linear, or total, order if for every $x, y \in X$ either $x \preceq y$ or $y \preceq x$, otherwise it is a partial order.

If \preceq is an order, then we also write $y \succeq x$ instead of $x \preceq y$, and we define a corresponding strict order \prec by

$$x \prec y \text{ if } x \preceq y \text{ and } x \neq y.$$

There are many ways to order a given set (with two or more elements).

Example 1.34. Let X be a set. One way to partially order the subsets of X is by inclusion, as in Example 1.32. Another way is to say that $A \preceq B$ for $A, B \subset X$ if and only if $A \supset B$, meaning that A is “smaller” than B if A includes B . Then \preceq is in an order on $\mathcal{P}(X)$, called ordering by reverse inclusion.

1.5.2. Equivalence relations. Equivalence relations decompose a set into disjoint subsets, called equivalence classes. We begin with an example of an equivalence relation on \mathbb{N} .

Example 1.35. Fix $N \in \mathbb{N}$ and say that $m \sim n$ if

$$m \equiv n \pmod{N},$$

meaning that $m - n$ is divisible by N . Two numbers are related by \sim if they have the same remainder when divided by N . Moreover, \mathbb{N} is the union of N equivalence classes, consisting of numbers with remainders $0, 1, \dots, N - 1$ modulo N .

The definition of an equivalence relation differs from the definition of an order only by changing antisymmetry to symmetry, but order relations and equivalence relations have completely different properties.

Definition 1.36. An equivalence relation \sim on a set X is a binary relation on X such that for every $x, y, z \in X$:

- (a) $x \sim x$ (reflexivity);
- (b) if $x \sim y$ then $y \sim x$ (symmetry);
- (c) if $x \sim y$ and $y \sim z$ then $x \sim z$ (transitivity).

For each $x \in X$, the set of elements equivalent to x ,

$$[x/\sim] = \{y \in X : x \sim y\},$$

is called the equivalence class of x with respect to \sim . When the equivalence relation is understood, we write the equivalence class $[x/\sim]$ simply as $[x]$. The set of equivalence classes of an equivalence relation \sim on a set X is denoted by X/\sim . Note that each element of X/\sim is a subset of X , so X/\sim is a subset of the power set $\mathcal{P}(X)$ of X .

The following theorem is the basic result about equivalence relations. It says that an equivalence relation on a set partitions the set into disjoint equivalence classes.

Theorem 1.37. Let \sim be an equivalence relation on a set X . Every equivalence class is non-empty, and X is the disjoint union of the equivalence classes of \sim .

Proof. If $x \in X$, then the symmetry of \sim implies that $x \in [x]$. Therefore every equivalence class is non-empty and the union of the equivalence classes is X .

To prove that the union is disjoint, we show that for every $x, y \in X$ either $[x] \cap [y] = \emptyset$ (if $x \not\sim y$) or $[x] = [y]$ (if $x \sim y$).

Suppose that $[x] \cap [y] \neq \emptyset$. Let $z \in [x] \cap [y]$ be an element in both equivalence classes. If $x_1 \in [x]$, then $x_1 \sim z$ and $z \sim y$, so $x_1 \sim y$ by the transitivity of \sim , and therefore $x_1 \in [y]$. It follows that $[x] \subset [y]$. A similar argument applied to $y_1 \in [y]$ implies that $[y] \subset [x]$, and therefore $[x] = [y]$. In particular, $y \in [x]$, so $x \sim y$. On the other hand, if $[x] \cap [y] = \emptyset$, then $y \notin [x]$ since $y \in [y]$, so $x \not\sim y$. \square

There is a natural projection $\pi : X \rightarrow X/\sim$, given by $\pi(x) = [x]$, that maps each element of X to the equivalence class that contains it. Conversely, we can index the collection of equivalence classes

$$X/\sim = \{[a] : a \in A\}$$

by a subset A of X which contains exactly one element from each equivalence class. It is important to recognize, however, that such an indexing involves an arbitrary choice of a representative element from each equivalence class, and it is better to think in terms of the collection of equivalence classes, rather than a subset of elements.

Example 1.38. The equivalence classes of \mathbb{N} relative to the equivalence relation $m \sim n$ if $m \equiv n \pmod{3}$ are given by

$$I_0 = \{3, 6, 9, \dots\}, \quad I_1 = \{1, 4, 7, \dots\}, \quad I_2 = \{2, 5, 8, \dots\}.$$

The projection $\pi : \mathbb{N} \rightarrow \{I_0, I_1, I_2\}$ maps a number to its equivalence class e.g. $\pi(101) = I_2$. We can choose $\{1, 2, 3\}$ as a set of representative elements, in which case

$$I_0 = [3], \quad I_1 = [1], \quad I_2 = [2],$$

but any other set $A \subset \mathbb{N}$ of three numbers with remainders 0, 1, 2 (mod 3) will do. For example, if we choose $A = \{7, 15, 101\}$, then

$$I_0 = [15], \quad I_1 = [7], \quad I_2 = [101].$$

1.6. Countable and uncountable sets

One way to show that two sets have the same “size” is to pair off their elements. For example, if we can match up every left shoe in a closet with a right shoe, with no right shoes left over, then we know that we have the same number of left and right shoes. That is, we have the same number of left and right shoes if there is a one-to-one, onto map $f : L \rightarrow R$, or one-to-one correspondence, from the set L of left shoes to the set R of right shoes.

We refer to the “size” of a set as measured by one-to-one correspondences as its cardinality. This notion enables us to compare the cardinality of both finite and infinite sets. In particular, we can use it to distinguish between “smaller” countably infinite sets, such as the integers or rational numbers, and “larger” uncountably infinite sets, such as the real numbers.

Definition 1.39. Two sets X, Y have equal cardinality, written $X \approx Y$, if there is a one-to-one, onto map $f : X \rightarrow Y$. The cardinality of X is less than or equal to the cardinality of Y , written $X \lesssim Y$, if there is a one-to-one (but not necessarily onto) map $g : X \rightarrow Y$.

If $X \approx Y$, then we also say that X, Y have the same cardinality. We don’t define the notion of a “cardinal number” here, only the relation between sets of “equal cardinality.”

Note that \approx is an equivalence relation on any collection of sets. In particular, it is transitive because if $X \approx Y$ and $Y \approx Z$, then there are one-to-one and onto maps $f : X \rightarrow Y$ and $g : Y \rightarrow Z$, so $g \circ f : X \rightarrow Z$ is one-to-one and onto, and $X \approx Z$. We may therefore divide any collection of sets into equivalence classes of sets with equal cardinality.

It follows immediately from the definition that \lesssim is reflexive and transitive. Furthermore, as stated in the following Schröder-Bernstein theorem, if $X \lesssim Y$ and $Y \lesssim X$, then $X \approx Y$. This result allows us to prove that two sets have equal cardinality by constructing one-to-one maps that need not be onto. The statement of the theorem is intuitively obvious but the proof, while elementary, is surprisingly involved and can be omitted without loss of continuity. (We will only use the theorem once, in the proof of Theorem 5.67.)

Theorem 1.40 (* Schröder-Bernstein). If X, Y are sets such that there are one-to-one maps $f : X \rightarrow Y$ and $g : Y \rightarrow X$, then there is a one-to-one, onto map $h : X \rightarrow Y$.

Proof. We divide X into three disjoint subsets X_X , X_Y , X_∞ with different mapping properties as follows.

Consider a point $x_1 \in X$. If x_1 is not in the range of g , then we say $x_1 \in X_X$. Otherwise there exists $y_1 \in Y$ such that $g(y_1) = x_1$, and y_1 is unique since g is one-to-one. If y_1 is not in the range of f , then we say $x_1 \in X_Y$. Otherwise there exists a unique $x_2 \in X$ such that $f(x_2) = y_1$. Continuing in this way, we generate a sequence of points

$$x_1, y_1, x_2, y_2, \dots, x_n, y_n, x_{n+1}, \dots$$

with $x_n \in X$, $y_n \in Y$ and

$$g(y_n) = x_n, \quad f(x_{n+1}) = y_n.$$

We assign the starting point x_1 to a subset in the following way: (a) $x_1 \in X_X$ if the sequence terminates at some $x_n \in X$ that isn't in the range of g ; (b) $x_1 \in X_Y$ if the sequence terminates at some $y_n \in Y$ that isn't in the range of f ; (c) $x_1 \in X_\infty$ if the sequence never terminates.

Similarly, if $y_1 \in Y$, then we generate a sequence of points

$$y_1, x_1, y_2, x_2, \dots, y_n, x_n, y_{n+1}, \dots$$

with $x_n \in X$, $y_n \in Y$ by

$$f(x_n) = y_n, \quad g(y_{n+1}) = x_n,$$

and we assign y_1 to a subset Y_X , Y_Y , or Y_∞ of Y as follows: (a) $y_1 \in Y_X$ if the sequence terminates at some $x_n \in X$ that isn't in the range of g ; (b) $y_1 \in Y_Y$ if the sequence terminates at some $y_n \in Y$ that isn't in the range of f ; (c) $y_1 \in Y_\infty$ if the sequence never terminates.

We claim that $f : X_X \rightarrow Y_X$ is one-to-one and onto. First, if $x \in X_X$, then $f(x) \in Y_X$ because the sequence generated by $f(x)$ coincides with the sequence generated by x after its first term, so both sequences terminate at a point in X . Second, if $y \in Y_X$, then there is $x \in X$ such that $f(x) = y$, otherwise the sequence would terminate at $y \in Y$, meaning that $y \in Y_Y$. Furthermore, we must have $x \in X_X$ because the sequence generated by x is a continuation of the sequence generated by y and therefore also terminates at a point in X . Finally, f is one-to-one on X_X since f is one-to-one on X .

The same argument applied to $g : Y_Y \rightarrow X_Y$ implies that g is one-to-one and onto, so $g^{-1} : X_Y \rightarrow Y_Y$ is one-to-one and onto.

Finally, similar arguments show that $f : X_\infty \rightarrow Y_\infty$ is one-to-one and onto: If $x \in X_\infty$, then the sequence generated by $f(x) \in Y$ doesn't terminate, so $f(x) \in Y_\infty$; and every $y \in Y_\infty$ is the image of a point $x \in X$ which, like y , generates a sequence that does not terminate, so $x \in X_\infty$.

It then follows that $h : X \rightarrow Y$ defined by

$$h(x) = \begin{cases} f(x) & \text{if } x \in X_X \\ g^{-1}(x) & \text{if } x \in X_Y \\ f(x) & \text{if } x \in X_\infty \end{cases}$$

is a one-to-one, onto map from X to Y . □

We can use the cardinality relation to describe the “size” of a set by comparing it with standard sets.

Definition 1.41. A set X is:

- (1) Finite if it is the empty set or $X \approx \{1, 2, \dots, n\}$ for some $n \in \mathbb{N}$;
- (2) Countably infinite (or denumerable) if $X \approx \mathbb{N}$;
- (3) Infinite if it is not finite;
- (4) Countable if it is finite or countably infinite;
- (5) Uncountable if it is not countable.

We’ll take for granted some intuitively obvious facts which follow from the definitions. For example, a finite, non-empty set is in one-to-one correspondence with $\{1, 2, \dots, n\}$ for a unique natural number $n \in \mathbb{N}$ (the number of elements in the set), a countably infinite set is not finite, and a subset of a countable set is countable.

According to Definition 1.41, we may divide sets into disjoint classes of finite, countably infinite, and uncountable sets. We also distinguish between finite and infinite sets, and countable and uncountable sets. We will show below, in Theorem 2.19, that the set of real numbers is uncountable, and we refer to its cardinality as the cardinality of the continuum.

Definition 1.42. A set X has the cardinality of the continuum if $X \approx \mathbb{R}$.

One has to be careful in extrapolating properties of finite sets to infinite sets.

Example 1.43. The set of squares

$$S = \{1, 4, 9, 16, \dots, n^2, \dots\}$$

is countably infinite since $f : \mathbb{N} \rightarrow S$ defined by $f(n) = n^2$ is one-to-one and onto. It may appear surprising at first that the set \mathbb{N} can be in one-to-one correspondence with an apparently “smaller” proper subset S , since this doesn’t happen for finite sets. In fact, assuming the axiom of choice, one can show that a set is infinite if and only if it has the same cardinality as a proper subset. Dedekind (1888) used this property to give a definition infinite sets that did not depend on the natural numbers \mathbb{N} .

Next, we prove some results about countable sets. The following proposition states a useful necessary and sufficient condition for a set to be countable.

Proposition 1.44. A non-empty set X is countable if and only if there is an onto map $f : \mathbb{N} \rightarrow X$.

Proof. If X is countably infinite, then there is a one-to-one, onto map $f : \mathbb{N} \rightarrow X$. If X is finite and non-empty, then for some $n \in \mathbb{N}$ there is a one-to-one, onto map $g : \{1, 2, \dots, n\} \rightarrow X$. Choose any $x \in X$ and define the onto map $f : \mathbb{N} \rightarrow X$ by

$$f(k) = \begin{cases} g(k) & \text{if } k = 1, 2, \dots, n, \\ x & \text{if } k = n + 1, n + 2, \dots \end{cases}$$

Conversely, suppose that such an onto map exists. We define a one-to-one, onto map g recursively by omitting repeated values of f . Explicitly, let $g(1) = f(1)$. Suppose that $n \geq 1$ and we have chosen n distinct g -values $g(1), g(2), \dots, g(n)$. Let

$$A_n = \{k \in \mathbb{N} : f(k) \neq g(j) \text{ for every } j = 1, 2, \dots, n\}$$

denote the set of natural numbers whose f -values are not already included among the g -values. If $A_n = \emptyset$, then $g : \{1, 2, \dots, n\} \rightarrow X$ is one-to-one and onto, and X is finite. Otherwise, let $k_n = \min A_n$, and define $g(n+1) = f(k_n)$, which is distinct from all of the previous g -values. Either this process terminates, and X is finite, or we go through all the f -values and obtain a one-to-one, onto map $g : \mathbb{N} \rightarrow X$, and X is countably infinite. \square

If X is a countable set, then we refer to an onto function $f : \mathbb{N} \rightarrow X$ as an enumeration of X , and write $X = \{x_n : n \in \mathbb{N}\}$, where $x_n = f(n)$.

Proposition 1.45. The Cartesian product $\mathbb{N} \times \mathbb{N}$ is countably infinite.

Proof. Define a linear order \prec on ordered pairs of natural numbers as follows:

$$(m, n) \prec (m', n') \quad \text{if either } m + n < m' + n' \text{ or } m + n = m' + n' \text{ and } n < n'.$$

That is, we arrange $\mathbb{N} \times \mathbb{N}$ in a table

$$\begin{array}{ccccc} (1, 1) & (1, 2) & (1, 3) & (1, 4) & \dots \\ (2, 1) & (2, 2) & (2, 3) & (2, 4) & \dots \\ (3, 1) & (3, 2) & (3, 3) & (3, 4) & \dots \\ (4, 1) & (4, 2) & (4, 3) & (4, 4) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

and list it along successive diagonals from bottom-left to top-right as

$$(1, 1), (2, 1), (1, 2), (3, 1), (2, 2), (1, 3), (4, 1), (3, 2), (2, 3), (1, 4), \dots$$

We define $f : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ by setting $f(n)$ equal to the n th pair in this order; for example, $f(7) = (4, 1)$. Then f is one-to-one and onto, so $\mathbb{N} \times \mathbb{N}$ is countably infinite. \square

Theorem 1.46. A countable union of countable sets is countable.

Proof. Let $\{X_n : n \in \mathbb{N}\}$ be a countable collection of countable sets. From Proposition 1.44, there is an onto map $f_n : \mathbb{N} \rightarrow X_n$. We define

$$g : \mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} X_n$$

by $g(n, k) = f_n(k)$. Then g is also onto. From Proposition 1.45, there is a one-to-one, onto map $h : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$, and it follows that

$$g \circ h : \mathbb{N} \rightarrow \bigcup_{n \in \mathbb{N}} X_n$$

is onto, so Proposition 1.44 implies that the union of the X_n is countable. \square

The next theorem gives a fundamental example of an uncountable set, namely the set of all subsets of natural numbers. The proof uses a “diagonal” argument due to Cantor (1891), which is of frequent use in analysis. Recall from Definition 1.1 that the power set of a set is the collection of all its subsets.

Theorem 1.47. The power set $\mathcal{P}(\mathbb{N})$ of \mathbb{N} is uncountable.

Proof. Let $\mathcal{C} \subset \mathcal{P}(\mathbb{N})$ be a countable collection of subsets of \mathbb{N}

$$\mathcal{C} = \{A_n \subset \mathbb{N} : n \in \mathbb{N}\}.$$

Define a subset $A \subset \mathbb{N}$ by

$$A = \{n \in \mathbb{N} : n \notin A_n\}.$$

Then $A \neq A_n$ for every $n \in \mathbb{N}$ since either $n \in A$ and $n \notin A_n$ or $n \notin A$ and $n \in A_n$. Thus, $A \notin \mathcal{C}$. It follows that no countable collection of subsets of \mathbb{N} includes all of the subsets of \mathbb{N} , so $\mathcal{P}(\mathbb{N})$ is uncountable. \square

This theorem has an immediate corollary for the set Σ of binary sequences defined in Definition 1.29.

Corollary 1.48. The set Σ of binary sequences has the same cardinality as $\mathcal{P}(\mathbb{N})$ and is uncountable.

Proof. By Example 1.30, the set Σ is in one-to-one correspondence with $\mathcal{P}(\mathbb{N})$, which is uncountable. \square

It is instructive to write the diagonal argument in terms of binary sequences. Suppose that $S = \{s_n \in \Sigma : n \in \mathbb{N}\}$ is a countable set of binary sequences that begins, for example, as follows

$$\begin{aligned} s_1 &= 001101 \dots \\ s_2 &= 110010 \dots \\ s_3 &= 110110 \dots \\ s_4 &= 011000 \dots \\ s_5 &= 100111 \dots \\ s_6 &= 100100 \dots \\ &\vdots \end{aligned}$$

Then we get a sequence $s \notin S$ by going down the diagonal and switching the values from 0 to 1 or from 1 to 0. For the previous sequences, this gives

$$s = 101101 \dots$$

We will show in Theorem 5.67 below that Σ and $\mathcal{P}(\mathbb{N})$ are also in one-to-one correspondence with \mathbb{R} , so both have the cardinality of the continuum.

A similar diagonal argument to the one used in Theorem 1.47 shows that for every set X the cardinality of the power set $\mathcal{P}(X)$ is strictly greater than the cardinality of X . In particular, the cardinality of $\mathcal{P}(\mathcal{P}(\mathbb{N}))$ is strictly greater than the cardinality of $\mathcal{P}(\mathbb{N})$, the cardinality of $\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))$ is strictly greater than

the cardinality of $\mathcal{P}(\mathcal{P}(\mathbb{N}))$, and so on. Thus, there are many other uncountable cardinalities apart from the cardinality of the continuum.

Cantor (1878) raised the question of whether or not there are any sets whose cardinality lies strictly between that of \mathbb{N} and $\mathcal{P}(\mathbb{N})$. The statement that there are no such sets is called the continuum hypothesis, which may be formulated as follows.

Hypothesis 1.49 (Continuum). If $\mathcal{C} \subset \mathcal{P}(\mathbb{N})$ is infinite, then either $\mathcal{C} \approx \mathbb{N}$ or $\mathcal{C} \approx \mathcal{P}(\mathbb{N})$.

The work of Gödel (1940) and Cohen (1963) established the remarkable result that the continuum hypothesis cannot be proved or disproved from the standard axioms of set theory (assuming, as we believe to be the case, that these axioms are consistent). This result illustrates a fundamental and unavoidable incompleteness in the ability of any finite system of axioms to capture the properties of any mathematical structure that is rich enough to include the natural numbers.

Chapter six:

Numbers and functions

The subject of this course is “functions of one real variable” so we begin by wondering what a real number “really” is, and then, in the next section, what a function is.

1. What is a number?

1.1. Different kinds of numbers. The simplest numbers are the *positive integers*

$$1, 2, 3, 4, \dots$$

the number *zero*

$$0,$$

and the *negative integers*

$$\dots, -4, -3, -2, -1.$$

Together these form the integers or “whole numbers.”

Next, there are the numbers you get by dividing one whole number by another (nonzero) whole number. These are the so called fractions or *rational numbers* such as

$$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{4}{3}, \dots$$

or

$$-\frac{1}{2}, -\frac{1}{3}, -\frac{2}{3}, -\frac{1}{4}, -\frac{2}{4}, -\frac{3}{4}, -\frac{4}{3}, \dots$$

By definition, any whole number is a rational number (in particular zero is a rational number.)

You can add, subtract, multiply and divide any pair of rational numbers and the result will again be a rational number (provided you don't try to divide by zero).

One day in middle school you were told that there are other numbers besides the rational numbers, and the first example of such a number is the square root of two. It has been known ever since the time of the greeks that no rational number exists whose square is exactly 2, i.e. you can't find a fraction $\frac{m}{n}$ such that

$$\left(\frac{m}{n}\right)^2 = 2, \text{ i.e. } m^2 = 2n^2.$$

Nevertheless, if you compute x^2 for some values of x between 1 and 2, and check if you get more or less than 2, then it looks like there should be some number x between 1.4 and 1.5 whose square is exactly 2. So, we *assume* that there is such a number, and we call it the square root of 2, written as $\sqrt{2}$. This raises several questions. How do we know there really is a number between 1.4 and 1.5 for which $x^2 = 2$? How many other such numbers are we going to assume into existence? Do these new numbers obey the same algebra rules (like $a + b = b + a$) as the rational numbers? If we knew precisely what these numbers (like $\sqrt{2}$) were then we could perhaps answer such questions. It turns out to be rather difficult to give a precise description of what a number is, and in this course we won't try to get anywhere near the bottom of this issue. Instead we will think of numbers as “infinite decimal expansions” as follows.

| x | x^2 |
|-----|------------|
| 1.2 | 1.44 |
| 1.3 | 1.69 |
| 1.4 | $1.96 < 2$ |
| 1.5 | $2.25 > 2$ |
| 1.6 | 2.56 |

One can represent certain fractions as decimal fractions, e.g.

$$\frac{279}{25} = \frac{1116}{100} = 11.16.$$

Not all fractions can be represented as decimal fractions. For instance, expanding $\frac{1}{3}$ into a decimal fraction leads to an unending decimal fraction

$$\frac{1}{3} = 0.333\ 333\ 333\ 333\ 333\ \dots$$

It is impossible to write the complete decimal expansion of $\frac{1}{3}$ because it contains infinitely many digits. But we can describe the expansion: each digit is a three. An electronic calculator, which always represents numbers as **finite** decimal numbers, can never hold the number $\frac{1}{3}$ exactly.

Every fraction can be written as a decimal fraction which may or may not be finite. If the decimal expansion doesn't end, then it must repeat. For instance,

$$\frac{1}{7} = 0.142857\ 142857\ 142857\ 142857\ \dots$$

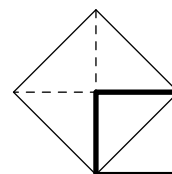
Conversely, any infinite repeating decimal expansion represents a rational number.

A **real number** is specified by a possibly unending decimal expansion. For instance,

$$\sqrt{2} = 1.414\ 213\ 562\ 373\ 095\ 048\ 801\ 688\ 724\ 209\ 698\ 078\ 569\ 671\ 875\ 376\ 9\ \dots$$

Of course you can never write *all* the digits in the decimal expansion, so you only write the first few digits and hide the others behind dots. To give a precise description of a real number (such as $\sqrt{2}$) you have to explain how you could *in principle* compute as many digits in the expansion as you would like. During the next three semesters of calculus we will not go into the details of how this should be done.

1.2. A reason to believe in $\sqrt{2}$. The Pythagorean theorem says that the hypotenuse of a right triangle with sides 1 and 1 must be a line segment of length $\sqrt{2}$. In middle or high school you learned something similar to the following geometric construction of a line segment whose length is $\sqrt{2}$. Take a square with side of length 1, and construct a new square one of whose sides is the diagonal of the first square. The figure you get consists of 5 triangles of equal area and by counting triangles you see that the larger square has exactly twice the area of the smaller square. Therefore the diagonal of the smaller square, being the side of the larger square, is $\sqrt{2}$ as long as the side of the smaller square.



Why are real numbers called real? All the numbers we will use in this first semester of calculus are “real numbers.” At some point (in 2nd semester calculus) it becomes useful to assume that there is a number whose square is -1 . No real number has this property since the square of any real number is positive, so it was decided to call this new imagined number “imaginary” and to refer to the numbers we already have (rationals, $\sqrt{2}$ -like things) as “real.”

1.3. The real number line and intervals. It is customary to visualize the real numbers as points on a straight line. We imagine a line, and choose one point on this line, which we call the **origin**. We also decide which direction we call “left” and hence which we call “right.” Some draw the number line vertically and use the words “up” and “down.”

To plot any real number x one marks off a distance x from the origin, to the right (up) if $x > 0$, to the left (down) if $x < 0$.

The **distance along the number line** between two numbers x and y is $|x - y|$. In particular, the distance is never a negative number.

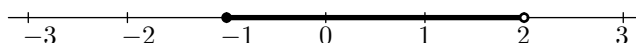


Figure 1. To draw the half open interval $[-1, 2)$ use a filled dot to mark the endpoint which is included and an open dot for an excluded endpoint.

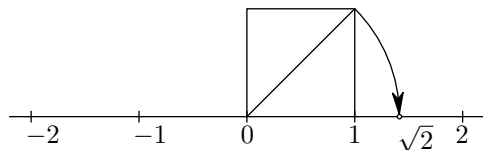


Figure 2. To find $\sqrt{2}$ on the real line you draw a square of sides 1 and drop the diagonal onto the real line.

Almost every equation involving variables x , y , etc. we write down in this course will be true for some values of x but not for others. In modern abstract mathematics a collection of real numbers (or any other kind of mathematical objects) is called a **set**. Below are some examples of sets of real numbers. We will use the notation from these examples throughout this course.

The collection of all real numbers between two given real numbers form an interval. The following notation is used

- (a, b) is the set of all real numbers x which satisfy $a < x < b$.
- $[a, b)$ is the set of all real numbers x which satisfy $a \leq x < b$.
- $(a, b]$ is the set of all real numbers x which satisfy $a < x \leq b$.
- $[a, b]$ is the set of all real numbers x which satisfy $a \leq x \leq b$.

If the endpoint is not included then it may be ∞ or $-\infty$. E.g. $(-\infty, 2]$ is the interval of all real numbers (both positive and negative) which are ≤ 2 .

1.4. Set notation. A common way of describing a set is to say it is the collection of all real numbers which satisfy a certain condition. One uses this notation

$$\mathcal{A} = \{x \mid x \text{ satisfies this or that condition}\}$$

Most of the time we will use upper case letters in a calligraphic font to denote sets. $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \dots)$

For instance, the interval (a, b) can be described as

$$(a, b) = \{x \mid a < x < b\}$$

The set

$$\mathcal{B} = \{x \mid x^2 - 1 > 0\}$$

consists of all real numbers x for which $x^2 - 1 > 0$, i.e. it consists of all real numbers x for which either $x > 1$ or $x < -1$ holds. This set consists of two parts: the interval $(-\infty, -1)$ and the interval $(1, \infty)$.

You can try to draw a set of real numbers by drawing the number line and coloring the points belonging to that set red, or by marking them in some other way.

Some sets can be very difficult to draw. For instance,

$$\mathcal{C} = \{x \mid x \text{ is a rational number}\}$$

can't be accurately drawn. In this course we will try to avoid such sets.

Sets can also contain just a few numbers, like

$$\mathcal{D} = \{1, 2, 3\}$$

which is the set containing the numbers one, two and three. Or the set

$$\mathcal{E} = \{x \mid x^3 - 4x^2 + 1 = 0\}$$

which consists of the solutions of the equation $x^3 - 4x^2 + 1 = 0$. (There are three of them, but it is not easy to give a formula for the solutions.)

If \mathcal{A} and \mathcal{B} are two sets then **the union of \mathcal{A} and \mathcal{B}** is the set which contains all numbers that belong either to \mathcal{A} or to \mathcal{B} . The following notation is used

$$\mathcal{A} \cup \mathcal{B} = \{x \mid x \text{ belongs to } \mathcal{A} \text{ or to } \mathcal{B} \text{ or both.}\}$$

Similarly, the *intersection of two sets \mathcal{A} and \mathcal{B}* is the set of numbers which belong to both sets. This notation is used:

$$\mathcal{A} \cap \mathcal{B} = \{x \mid x \text{ belongs to both } \mathcal{A} \text{ and } \mathcal{B}\}$$

2. Exercises

1. What is the 2007th digit after the period in the expansion of $\frac{1}{7}$?

2. Which of the following fractions have finite decimal expansions?

$$a = \frac{2}{3}, \quad b = \frac{3}{25}, \quad c = \frac{276937}{15625}.$$

3. Draw the following sets of real numbers. Each of these sets is the union of one or more intervals. Find those intervals. Which of these sets are finite?

$$\mathcal{A} = \{x \mid x^2 - 3x + 2 \leq 0\}$$

$$\mathcal{B} = \{x \mid x^2 - 3x + 2 \geq 0\}$$

$$\mathcal{C} = \{x \mid x^2 - 3x > 3\}$$

$$\mathcal{D} = \{x \mid x^2 - 5 > 2x\}$$

$$\mathcal{E} = \{t \mid t^2 - 3t + 2 \leq 0\}$$

$$\mathcal{F} = \{\alpha \mid \alpha^2 - 3\alpha + 2 \geq 0\}$$

$$\mathcal{G} = (0, 1) \cup (5, 7]$$

$$\mathcal{H} = (\{1\} \cup \{2, 3\}) \cap (0, 2\sqrt{2})$$

$$\mathcal{Q} = \{\theta \mid \sin \theta = \frac{1}{2}\}$$

$$\mathcal{R} = \{\varphi \mid \cos \varphi > 0\}$$

4. Suppose \mathcal{A} and \mathcal{B} are intervals. Is it always true that $\mathcal{A} \cap \mathcal{B}$ is an interval? How about $\mathcal{A} \cup \mathcal{B}$?

5. Consider the sets

$$\mathcal{M} = \{x \mid x > 0\} \text{ and } \mathcal{N} = \{y \mid y > 0\}.$$

Are these sets the same?

6. **Group Problem.**

Write the numbers

$$x = 0.3131313131\dots, \quad y = 0.273273273273\dots$$

$$\text{and } z = 0.21541541541541541\dots$$

as fractions (i.e. write them as $\frac{m}{n}$, specifying m and n .)

(Hint: show that $100x = x + 31$. A similar trick works for y , but z is a little harder.)

7. **Group Problem.**

Is the number whose decimal expansion after the period consists only of nines, i.e.

$$x = 0.9999999999999999\dots$$

an integer?

3. Functions

Wherein we meet the main characters of this semester

3.1. Definition. To specify a *function* f you must

- (1) give a *rule* which tells you how to compute the value $f(x)$ of the function for a given real number x , and:
- (2) say for which real numbers x the rule may be applied.

The set of numbers for which a function is defined is called its *domain*. The set of all possible numbers $f(x)$ as x runs over the domain is called the *range* of the function. The rule must be *unambiguous*: the same x must always lead to the same $f(x)$.

For instance, one can define a function f by putting $f(x) = \sqrt{x}$ for all $x \geq 0$. Here the rule defining f is “take the square root of whatever number you’re given”, and the function f will accept all nonnegative real numbers.

The rule which specifies a function can come in many different forms. Most often it is a formula, as in the square root example of the previous paragraph. Sometimes you need a few formulas, as in

$$g(x) = \begin{cases} 2x & \text{for } x < 0 \\ x^2 & \text{for } x \geq 0 \end{cases} \quad \text{domain of } g = \text{all real numbers.}$$

Functions which are defined by different formulas on different intervals are sometimes called *piecewise defined functions*.

3.2. Graphing a function. You get the *graph of a function* f by drawing all points whose coordinates are (x, y) where x must be in the domain of f and $y = f(x)$.

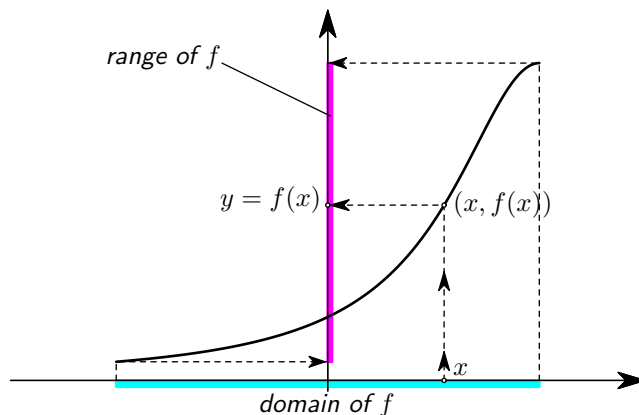


Figure 3. The graph of a function f . The domain of f consists of all x values at which the function is defined, and the range consists of all possible values f can have.

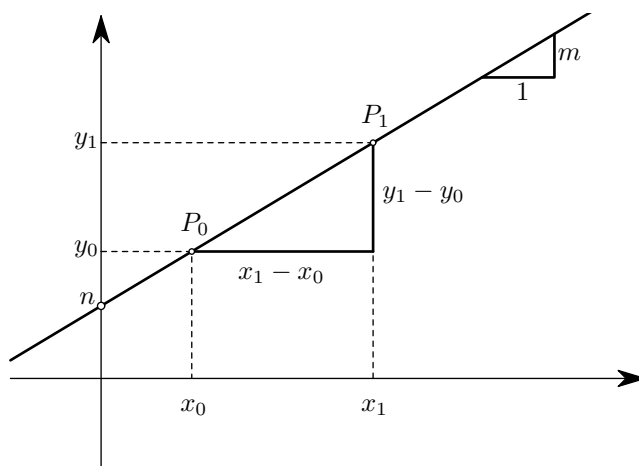


Figure 4. A straight line and its slope. The line is the graph of $f(x) = mx + n$. It intersects the y -axis at height n , and the ratio between the amounts by which y and x increase as you move from one point to another on the line is $\frac{y_1 - y_0}{x_1 - x_0} = m$.

3.3. Linear functions. A function which is given by the formula

$$f(x) = mx + n$$

where m and n are constants is called a **linear function**. Its graph is a straight line. The constants m and n are the **slope** and **y -intercept** of the line. Conversely, any straight line which is not vertical (i.e. not parallel to the y -axis) is the graph of a linear function. If you know two points (x_0, y_0) and (x_1, y_1) on the line, then then one can compute the slope m from the “rise-over-run” formula

$$m = \frac{y_1 - y_0}{x_1 - x_0}.$$

This formula actually contains a theorem from Euclidean geometry, namely it says that the ratio $(y_1 - y_0) : (x_1 - x_0)$ is the same for every pair of points (x_0, y_0) and (x_1, y_1) that you could pick on the line.

3.4. Domain and “biggest possible domain.” In this course we will usually not be careful about specifying the domain of the function. When this happens the domain is understood to be the set of all x for which the rule which tells you how to compute $f(x)$ is meaningful. For instance, if we say that h is the function

$$h(x) = \sqrt{x}$$

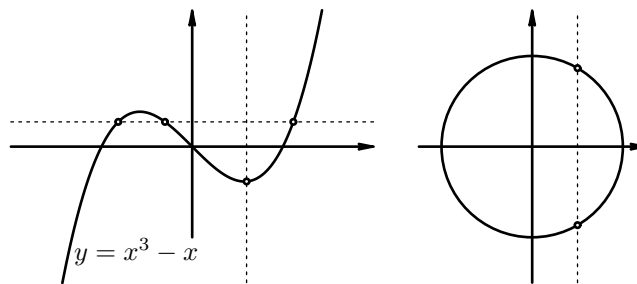


Figure 5. The graph of $y = x^3 - x$ fails the “horizontal line test,” but it passes the “vertical line test.” The circle fails both tests.

then the domain of h is understood to be the set of all nonnegative real numbers

$$\text{domain of } h = [0, \infty)$$

since \sqrt{x} is well-defined for all $x \geq 0$ and undefined for $x < 0$.

A systematic way of finding the domain and range of a function for which you are only given a formula is as follows:

- The domain of f consists of all x for which $f(x)$ is well-defined (“makes sense”)
- The range of f consists of all y for which you can solve the equation $f(x) = y$.

3.5. Example – find the domain and range of $f(x) = 1/x^2$. The expression $1/x^2$ can be computed for all real numbers x except $x = 0$ since this leads to division by zero. Hence the domain of the function $f(x) = 1/x^2$ is

$$\text{“all real numbers except 0”} = \{x \mid x \neq 0\} = (-\infty, 0) \cup (0, \infty).$$

To find the range we ask “for which y can we solve the equation $y = f(x)$ for x ,” i.e. we for which y can you solve $y = 1/x^2$ for x ?

If $y = 1/x^2$ then we must have $x^2 = 1/y$, so first of all, since we have to divide by y , y can’t be zero. Furthermore, $1/y = x^2$ says that y must be positive. On the other hand, if $y > 0$ then $y = 1/x^2$ has a solution (in fact two solutions), namely $x = \pm 1/\sqrt{y}$. This shows that the range of f is

$$\text{“all positive real numbers”} = \{x \mid x > 0\} = (0, \infty).$$

3.6. Functions in “real life.” One can describe the motion of an object using a function. If some object is moving along a straight line, then you can define the following function: Let $x(t)$ be the distance from the object to a fixed marker on the line, at the time t . Here the domain of the function is the set of all times t for which we know the position of the object, and the rule is

Given t , measure the distance between the object and the marker at time t .

There are many examples of this kind. For instance, a biologist could describe the growth of a cell by defining $m(t)$ to be the mass of the cell at time t (measured since the birth of the cell). Here the domain is the interval $[0, T]$, where T is the life time of the cell, and the rule that describes the function is

Given t , weigh the cell at time t .

3.7. The Vertical Line Property. Generally speaking graphs of functions are curves in the plane but they distinguish themselves from arbitrary curves by the way they intersect vertical lines: **The graph of a function cannot intersect a vertical line “ $x = \text{constant}$ ” in more than one point.** The reason why this is true is very simple: if two points lie on a vertical line, then they have the same x coordinate, so if they also lie on the graph of a function f , then their y -coordinates must also be equal, namely $f(x)$.

3.8. Examples. The graph of $f(x) = x^3 - x$ “goes up and down,” and, even though it intersects several horizontal lines in more than one point, it intersects *every* vertical line in exactly one point.

The collection of points determined by the equation $x^2 + y^2 = 1$ is a circle. It is not the graph of a function since the vertical line $x = 0$ (the y -axis) intersects the graph in two points $P_1(0, 1)$ and $P_2(0, -1)$. See Figure 6.

4. Inverse functions and Implicit functions

For many functions the rule which tells you how to compute it is not an explicit formula, but instead an equation which you still must solve. A function which is defined in this way is called an “implicit function.”

4.1. Example. One can define a function f by saying that for each x the value of $f(x)$ is the solution y of the equation

$$x^2 + 2y - 3 = 0.$$

In this example you can solve the equation for y ,

$$y = \frac{3 - x^2}{2}.$$

Thus we see that the function we have defined is $f(x) = (3 - x^2)/2$.

Here we have two definitions of the same function, namely

- (i) “ $y = f(x)$ is defined by $x^2 + 2y - 3 = 0$,” and
- (ii) “ f is defined by $f(x) = (3 - x^2)/2$.”

The first definition is the implicit definition, the second is explicit. You see that with an “implicit function” it isn’t the function itself, but rather the way it was defined that’s implicit.

4.2. Another example: domain of an implicitly defined function. Define g by saying that for any x the value $y = g(x)$ is the solution of

$$x^2 + xy - 3 = 0.$$

Just as in the previous example one can then solve for y , and one finds that

$$g(x) = y = \frac{3 - x^2}{x}.$$

Unlike the previous example this formula does not make sense when $x = 0$, and indeed, for $x = 0$ our rule for g says that $g(0) = y$ is the solution of

$$0^2 + 0 \cdot y - 3 = 0, \text{ i.e. } y \text{ is the solution of } 3 = 0.$$

That equation has no solution and hence $x = 0$ does not belong to the domain of our function g .

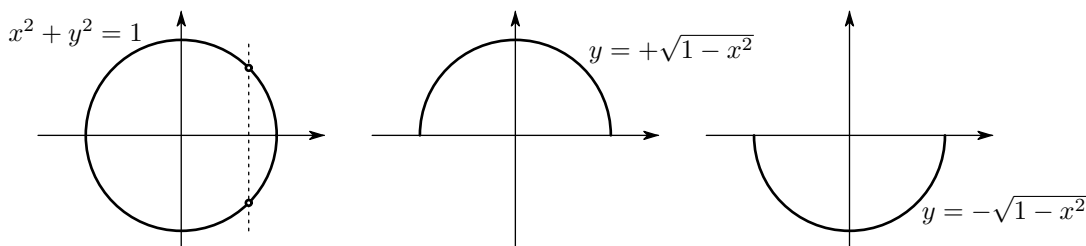


Figure 6. The circle determined by $x^2 + y^2 = 1$ is not the graph of a function, but it contains the graphs of the two functions $h_1(x) = \sqrt{1 - x^2}$ and $h_2(x) = -\sqrt{1 - x^2}$.

4.3. Example: the equation alone does not determine the function. Define $y = h(x)$ to be the solution of

$$x^2 + y^2 = 1.$$

If $x > 1$ or $x < -1$ then $x^2 > 1$ and there is no solution, so $h(x)$ is at most defined when $-1 \leq x \leq 1$. But when $-1 < x < 1$ there is another problem: not only does the equation have a solution, but it even has two solutions:

$$x^2 + y^2 = 1 \iff y = \sqrt{1 - x^2} \text{ or } y = -\sqrt{1 - x^2}.$$

The rule which defines a function must be unambiguous, and since we have not specified which of these two solutions is $h(x)$ the function is not defined for $-1 < x < 1$.

One can fix this by making a choice, but there are many possible choices. Here are three possibilities:

$$h_1(x) = \text{the nonnegative solution } y \text{ of } x^2 + y^2 = 1$$

$$h_2(x) = \text{the nonpositive solution } y \text{ of } x^2 + y^2 = 1$$

$$h_3(x) = \begin{cases} h_1(x) & \text{when } x < 0 \\ h_2(x) & \text{when } x \geq 0 \end{cases}$$

4.4. Why use implicit functions? In all the examples we have done so far we could replace the implicit description of the function with an explicit formula. This is not always possible or if it is possible the implicit description is much simpler than the explicit formula. For instance, you can define a function f by saying that $y = f(x)$ if and only if

$$(1) \quad y^3 + 3y + 2x = 0.$$

This means that the recipe for computing $f(x)$ for any given x is “solve the equation $y^3 + 3y + 2x = 0$.” E.g. to compute $f(0)$ you set $x = 0$ and solve $y^3 + 3y = 0$. The only solution is $y = 0$, so $f(0) = 0$. To compute $f(1)$ you have to solve $y^3 + 3y + 2 \cdot 1 = 0$, and if you’re lucky you see that $y = -1$ is the solution, and $f(1) = -1$.

In general, no matter what x is, the equation (1) turns out to have exactly one solution y (which depends on x , this is how you get the function f). Solving (1) is not easy. In the early 1500s Cardano and Tartaglia discovered a formula¹ for the solution. Here it is:

$$y = f(x) = \sqrt[3]{-x + \sqrt{1 + x^2}} - \sqrt[3]{x + \sqrt{1 + x^2}}.$$

The implicit description looks a lot simpler, and when we try to differentiate this function later on, it will be much easier to use “implicit differentiation” than to use the Cardano-Tartaglia formula directly.

4.5. Inverse functions. If you have a function f , then you can try to define a new function f^{-1} , the so-called **inverse function of f** , by the following prescription:

$$(2) \quad \text{For any given } x \text{ we say that } y = f^{-1}(x) \text{ if } y \text{ is the solution to the equation } f(y) = x.$$

So to find $y = f^{-1}(x)$ you solve the equation $x = f(y)$. If this is to define a function then the prescription (2) must be unambiguous and the equation $f(y) = x$ has to have a solution and cannot have more than one solution.

¹To see the solution and its history visit

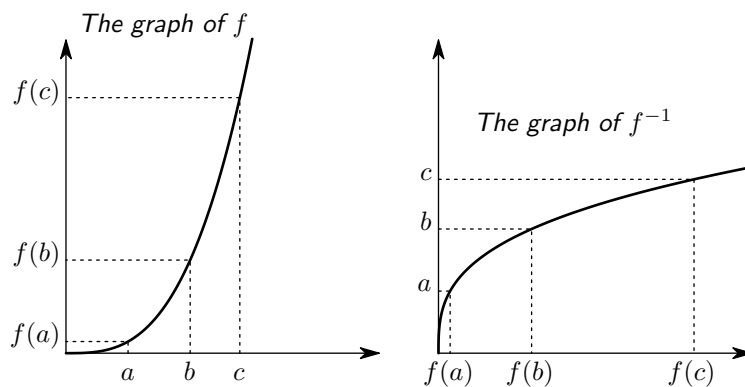


Figure 7. The graph of a function and its inverse are mirror images of each other.

4.6. Examples. Consider the function f with $f(x) = 2x + 3$. Then the equation $f(y) = x$ works out to be

$$2y + 3 = x$$

and this has the solution

$$y = \frac{x - 3}{2}.$$

So $f^{-1}(x)$ is defined for all x , and it is given by $f^{-1}(x) = (x - 3)/2$.

Next we consider the function $g(x) = x^2$ with domain all positive real numbers. To see for which x the inverse $g^{-1}(x)$ is defined we try to solve the equation $g(y) = x$, i.e. we try to solve $y^2 = x$. If $x < 0$ then this equation has no solutions since $y \geq 0$ for all y . But if $x \geq 0$ then $y = \sqrt{x}$ does have a solution, namely $y = \sqrt{x}$.

So we see that $g^{-1}(x)$ is defined for all nonnegative real numbers x , and that it is given by $g^{-1}(x) = \sqrt{x}$.

4.7. Inverse trigonometric functions. The familiar trigonometric functions Sine, Cosine and Tangent have inverses which are called arcsine, arccosine and arctangent.

| | | | |
|--------------|------------------------------|------------------|----------------------|
| $y = f(x)$ | | $x = f^{-1}(y)$ | |
| $y = \sin x$ | $(-\pi/2 \leq x \leq \pi/2)$ | $x = \arcsin(y)$ | $(-1 \leq y \leq 1)$ |
| $y = \cos x$ | $(0 \leq x \leq \pi)$ | $x = \arccos(y)$ | $(-1 \leq y \leq 1)$ |
| $y = \tan x$ | $(-\pi/2 < x < \pi/2)$ | $x = \arctan(y)$ | |

The notations $\arcsin y = \sin^{-1} y$, $\arccos x = \cos^{-1} x$, and $\arctan u = \tan^{-1} u$ are also commonly used for the inverse trigonometric functions. We will avoid the $\sin^{-1} y$ notation because it is ambiguous. Namely, everybody writes the square of $\sin y$ as

$$(\sin y)^2 = \sin^2 y.$$

Replacing the 2's by -1 's would lead to

$$\arcsin y = \sin^{-1} y \stackrel{??}{=} (\sin y)^{-1} = \frac{1}{\sin y}, \quad \text{which is not true!}$$

5. Exercises

- 8.** The functions f and g are defined by

$$f(x) = x^2 \text{ and } g(s) = s^2.$$

Are f and g the same functions or are they different?

- 9.** Find a formula for the function f which is defined by

$$y = f(x) \iff x^2 y + y = 7.$$

What is the domain of f ?

- 10.** Find a formula for the function f which is defined by

$$y = f(x) \iff x^2 y - y = 6.$$

What is the domain of f ?

- 11.** Let f be the function defined by $y = f(x) \iff y$ is the largest solution of

$$y^2 = 3x^2 - 2xy.$$

Find a formula for f . What are the domain and range of f ?

12. Find a formula for the function f which is defined by

$$y = f(x) \iff 2x + 2xy + y^2 = 5 \text{ and } y > -x.$$

Find the domain of f .

13. Use a calculator to compute $f(1.2)$ in three decimals where f is the implicitly defined function from §4.4. (There are (at least) two different ways of finding $f(1.2)$)

14. **Group Problem.**

(a) True or false:

for all x one has $\sin(\arcsin x) = x$?

(b) True or false:

for all x one has $\arcsin(\sin x) = x$?

15. On a graphing calculator plot the graphs of the following functions, and explain the results. (Hint: first do the previous exercise.)

$$f(x) = \arcsin(\sin x), \quad -2\pi \leq x \leq 2\pi$$

$$g(x) = \arcsin(x) + \arccos(x), \quad 0 \leq x \leq 1$$

$$h(x) = \arctan \frac{\sin x}{\cos x}, \quad |x| < \pi/2$$

$$k(x) = \arctan \frac{\cos x}{\sin x}, \quad |x| < \pi/2$$

$$l(x) = \arcsin(\cos x), \quad -\pi \leq x \leq \pi$$

$$m(x) = \cos(\arcsin x), \quad -1 \leq x \leq 1$$

16. Find the inverse of the function f which is given by $f(x) = \sin x$ and **whose domain is** $\pi \leq x \leq 2\pi$. Sketch the graphs of both f and f^{-1} .

17. Find a number a such that the function $f(x) = \sin(x + \pi/4)$ with domain $a \leq x \leq a + \pi$ has an inverse. Give a formula for $f^{-1}(x)$ using the arcsine function.

18. Draw the graph of the function h_3 from §4.3.

19. A function f is given which satisfies

$$f(2x + 3) = x^2$$

for all real numbers x .

Compute

$$(a) f(0) \quad (b) f(3) \quad (c) f(x)$$

$$(d) f(y) \quad (e) f(f(2))$$

where x and y are arbitrary real numbers.

What are the range and domain of f ?

20. A function f is given which satisfies

$$f\left(\frac{1}{x+1}\right) = 2x - 12.$$

for all real numbers x .

Compute

$$(a) f(1) \quad (b) f(0) \quad (c) f(x)$$

$$(d) f(t) \quad (e) f(f(2))$$

where x and t are arbitrary real numbers.

What are the range and domain of f ?

21. Does there exist a function f which satisfies

$$f(x^2) = x + 1$$

for **all** real numbers x ?

* * *

The following exercises review precalculus material involving quadratic expressions $ax^2 + bx + c$ in one way or another.

22. Explain how you "complete the square" in a quadratic expression like $ax^2 + bx$.

23. Find the range of the following functions:

$$f(x) = 2x^2 + 3$$

$$g(x) = -2x^2 + 4x$$

$$h(x) = 4x + x^2$$

$$k(x) = 4\sin x + \sin^2 x$$

$$\ell(x) = 1/(1 + x^2)$$

$$m(x) = 1/(3 + 2x + x^2).$$

24. **Group Problem.**

For each real number a we define a line ℓ_a with equation $y = ax + a^2$.

(a) Draw the lines corresponding to $a = -2, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2$.

(b) Does the point with coordinates $(3, 2)$ lie on one or more of the lines ℓ_a (where a can be any number, not just the five values from part (a))? If so, for which values of a does $(3, 2)$ lie on ℓ_a ?

(c) Which points in the plane lie on at least one of the lines ℓ_a ?

25. For which values of m and n does the graph of $f(x) = mx + n$ intersect the graph of $g(x) = 1/x$ in exactly one point and also contain the point $(-1, 1)$?

26. For which values of m and n does the graph of $f(x) = mx + n$ **not** intersect the graph of $g(x) = 1/x$?

Derivatives (1)

To work with derivatives you have to know what a limit is, but to motivate why we are going to study limits let's first look at the two classical problems that gave rise to the notion of a derivative: the tangent to a curve, and the instantaneous velocity of a moving object.

1. The tangent to a curve

Suppose you have a function $y = f(x)$ and you draw its graph. If you want to find the tangent to the graph of f at some given point on the graph of f , how would you do that?

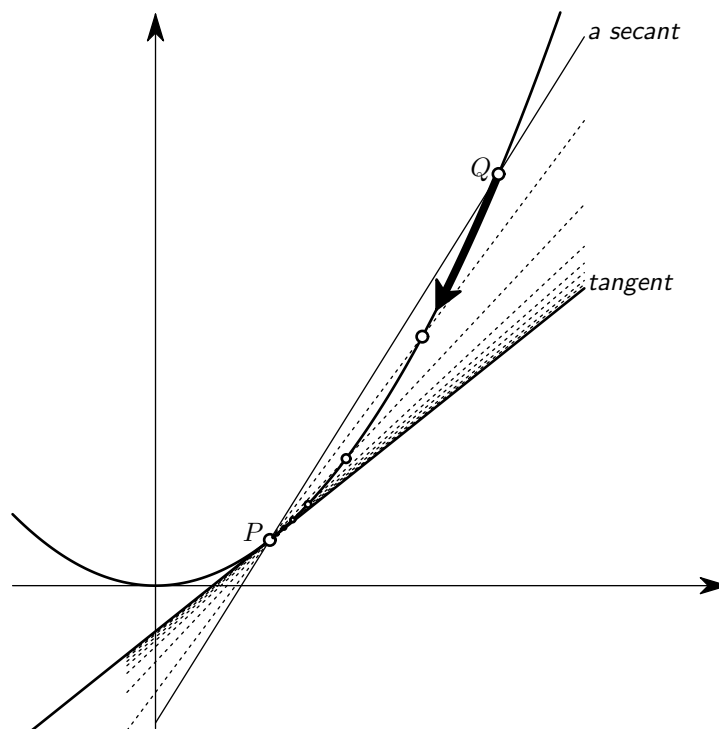


Figure 1. Constructing the tangent by letting $Q \rightarrow P$

Let P be the point on the graph at which you want to draw the tangent. If you are making a real paper and ink drawing you would take a ruler, make sure it goes through P and then turn it until it doesn't cross the graph anywhere else.

If you are using equations to describe the curve and lines, then you could pick a point Q on the graph and construct the line through P and Q ("construct" means "find an equation for"). This line is called a "secant," and it is of course not the tangent that you're looking for. But if you choose Q to be very close to P then the secant will be close to the tangent.

So this is our recipe for constructing the tangent through P : pick another point Q on the graph, find the line through P and Q , and see what happens to this line as you take Q closer and closer to P . The resulting secants will then get closer and closer to some line, and that line is the tangent.

We'll write this in formulas in a moment, but first let's worry about how close Q should be to P . We can't set Q equal to P , because then P and Q don't determine a line (you need **two** points to determine a line). If you choose Q different from P then you don't get the tangent, but at best something that is "close" to it. Some people have suggested that one should take Q "infinitely close" to P , but it isn't clear what that would mean. The concept of a limit is meant to solve this confusing problem.

2. An example – tangent to a parabola

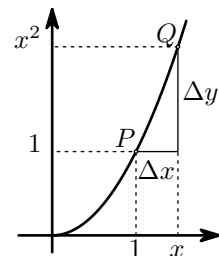
To make things more concrete, suppose that the function we had was $f(x) = x^2$, and that the point was $(1, 1)$. The graph of f is of course a parabola.

Any line through the point $P(1, 1)$ has equation

$$y - 1 = m(x - 1)$$

where m is the slope of the line. So instead of finding the equation of the secant and tangent lines we will find their slopes.

Let Q be the other point on the parabola, with coordinates (x, x^2) . We can "move Q around on the graph" by changing x . Whatever x we choose, it must be different from 1, for otherwise P and Q would be the same point. What we want to find out is how the line through P and Q changes if x is changed (and in particular, if x is chosen very close to a). Now, as one changes x one thing stays the same, namely, the secant still goes through P . So to describe the secant we only need to know its slope. By the "rise over run" formula, the slope of the secant line joining P and Q is



$$m_{PQ} = \frac{\Delta y}{\Delta x} \quad \text{where} \quad \Delta y = x^2 - 1 \quad \text{and} \quad \Delta x = x - 1.$$

By factoring $x^2 - 1$ we can rewrite the formula for the slope as follows

$$(3) \quad m_{PQ} = \frac{\Delta y}{\Delta x} = \frac{x^2 - 1}{x - 1} = \frac{(x - 1)(x + 1)}{x - 1} = x + 1.$$

As x gets closer to 1, the slope m_{PQ} , being $x + 1$, gets closer to the value $1 + 1 = 2$. We say that

the limit of the slope m_{PQ} as Q approaches P is 2.

In symbols,

$$\lim_{Q \rightarrow P} m_{PQ} = 2,$$

or, since Q approaching P is the same as x approaching 1,

$$(4) \quad \lim_{x \rightarrow 1} m_{PQ} = 2.$$

So we find that the tangent line to the parabola $y = x^2$ at the point $(1, 1)$ has equation

$$y - 1 = 2(x - 1), \text{ i.e. } y = 2x - 1.$$

A warning: you cannot substitute $x = 1$ in equation (3) to get (4) even though it looks like that's what we did. The reason why you can't do that is that when $x = 1$ the point Q coincides with the point P so "the line through P and Q " is not defined; also, if $x = 1$ then $\Delta x = \Delta y = 0$ so that the rise-over-run formula for the slope gives

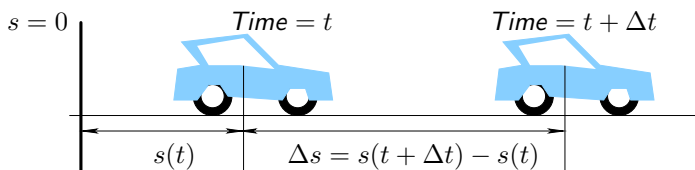
$$m_{PQ} = \frac{\Delta x}{\Delta y} = \frac{0}{0} = \text{undefined}.$$

It is only after the algebra trick in (3) that setting $x = 1$ gives something that is well defined. But if the intermediate steps leading to $m_{PQ} = x + 1$ aren't valid for $x = 1$ why should the final result mean anything for $x = 1$?

Something more complicated has happened. We did a calculation which is valid for all $x \neq 1$, and later looked at what happens if x gets “very close to 1.” This is the concept of a limit and we’ll study it in more detail later in this section, but first another example.

3. Instantaneous velocity

If you try to define “instantaneous velocity” you will again end up trying to divide zero by zero. Here is how it goes: When you are driving in your car the speedometer tells you how fast you are going, i.e. what your velocity is. What is this velocity? What does it mean if the speedometer says “50mph”?



We all know what **average velocity** is. Namely, if it takes you two hours to cover 100 miles, then your average velocity was

$$\frac{\text{distance traveled}}{\text{time it took}} = 50 \text{ miles per hour.}$$

This is not the number the speedometer provides you – it doesn’t wait two hours, measure how far you went and compute distance/time. If the speedometer in your car tells you that you are driving 50mph, then that should be your velocity **at the moment** that you look at your speedometer, i.e. “distance traveled over time it took” at the moment you look at the speedometer. But during the moment you look at your speedometer no time goes by (because a moment has no length) and you didn’t cover any distance, so your velocity at that moment is $\frac{0}{0}$, i.e. undefined. Your velocity at **any** moment is undefined. But then what is the speedometer telling you?

To put all this into formulas we need to introduce some notation. Let t be the time (in hours) that has passed since we got onto the road, and let $s(t)$ be the distance we have covered since then.

Instead of trying to find the velocity exactly at time t , we find a formula for the average velocity during some (short) time interval beginning at time t . We’ll write Δt for the length of the time interval.

At time t we have traveled $s(t)$ miles. A little later, at time $t + \Delta t$ we have traveled $s(t + \Delta t)$. Therefore during the time interval from t to $t + \Delta t$ we have moved $s(t + \Delta t) - s(t)$ miles. Our average velocity in that time interval is therefore

$$\frac{s(t + \Delta t) - s(t)}{\Delta t} \text{ miles per hour.}$$

The shorter you make the time interval, i.e. the smaller you choose Δt , the closer this number should be to the instantaneous velocity at time t .

So we have the following formula (definition, really) for the velocity at time t

$$(5) \quad v(t) = \lim_{\Delta t \rightarrow 0} \frac{s(t + \Delta t) - s(t)}{\Delta t}.$$

4. Rates of change

The two previous examples have much in common. If we ignore all the details about geometry, graphs, highways and motion, the following happened in both examples:

We had a function $y = f(x)$, and we wanted to know how much $f(x)$ changes if x changes. If you change x to $x + \Delta x$, then y will change from $f(x)$ to $f(x + \Delta x)$. The change in y is therefore

$$\Delta y = f(x + \Delta x) - f(x),$$

and the average rate of change is

$$(6) \quad \frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

This is the average rate of change of f over the interval from x to $x + \Delta x$. To define **the rate of change of the function f at x** we let the length Δx of the interval become smaller and smaller, in the hope that the average rate of change over the shorter and shorter time intervals will get closer and closer to some number. If that happens then that “limiting number” is called the rate of change of f at x , or, the **derivative** of f at x . It is written as

$$(7) \quad f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Derivatives and what you can do with them are what the first half of this semester is about. The description we just went through shows that to understand what a derivative is you need to know what a limit is. In the next chapter we'll study limits so that we get a less vague understanding of formulas like (7).

5. Examples of rates of change

5.1. Acceleration as the rate at which velocity changes. As you are driving in your car your velocity does not stay constant, it changes with time. Suppose $v(t)$ is your velocity at time t (measured in miles per hour). You could try to figure out how fast your velocity is changing by measuring it at one moment in time (you get $v(t)$), then measuring it a little later (you get $v(t + \Delta t)$). You conclude that your velocity increased by $\Delta v = v(t + \Delta t) - v(t)$ during a time interval of length Δt , and hence

$$\left\{ \begin{array}{l} \text{average rate at which} \\ \text{your velocity changed} \end{array} \right\} = \frac{\Delta v}{\Delta t} = \frac{v(t + \Delta t) - v(t)}{\Delta t}.$$

This rate of change is called your *average acceleration* (over the time interval from t to $t + \Delta t$). Your *instantaneous acceleration* at time t is the limit of your average acceleration as you make the time interval shorter and shorter:

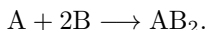
$$\{\text{acceleration at time } t\} = a = \lim_{\Delta t \rightarrow 0} \frac{v(t + \Delta t) - v(t)}{\Delta t}.$$

th the average and instantaneous accelerations are measured in “miles per hour per hour,” i.e. in

$$(\text{mi/h})/\text{h} = \text{mi/h}^2.$$

Or, if you had measured distances in meters and time in seconds then velocities would be measured in meters per second, and acceleration in meters per second per second, which is the same as meters per second², i.e. “meters per squared second.”

5.2. Reaction rates. Think of a chemical reaction in which two substances A and B react to form AB₂ according to the reaction



If the reaction is taking place in a closed reactor, then the “amounts” of A and B will be decreasing, while the amount of AB₂ will increase. Chemists write $[A]$ for the amount of “A” in the chemical reactor (measured in moles). Clearly $[A]$ changes with time so it defines a *function*. We’re mathematicians so we will write “ $[A](t)$ ” for the number of moles of A present at time t .

To describe how fast the amount of A is changing we consider the derivative of $[A]$ with respect to time, i.e.

$$[A]'(t) = \lim_{\Delta t \rightarrow 0} \frac{[A](t + \Delta t) - [A](t)}{\Delta t}.$$

This quantity is the rate of change of $[A]$. The notation “ $[A]'(t)$ ” is really only used by calculus professors. If you open a paper on chemistry you will find that the derivative is written in LEIBNIZ notation:

$$\frac{d[A]}{dt}$$

More on this in §1.2

How fast does the reaction take place? If you add more A or more B to the reactor then you would expect that the reaction would go faster, i.e. that more AB₂ is being produced per second. The law of *mass-action*

kinetics from chemistry states this more precisely. For our particular reaction it would say that the rate at which A is consumed is given by

$$\frac{d[A]}{dt} = k[A][B]^2,$$

in which the constant k is called the *reaction constant*. It's a constant that you could try to measure by timing how fast the reaction goes.

6. Exercises

27. Repeat the reasoning in §2 to find the slope at the point $(\frac{1}{2}, \frac{1}{4})$, or more generally at any point (a, a^2) on the parabola with equation $y = x^2$.

28. Repeat the reasoning in §2 to find the slope at the point $(\frac{1}{2}, \frac{1}{8})$, or more generally at any point (a, a^3) on the curve with equation $y = x^3$.

29. Group Problem.

Should you trust your calculator?

Find the slope of the tangent to the parabola $y = x^2$ at the point $(\frac{1}{3}, \frac{1}{9})$ (You have already done this: see exercise 27).

Instead of doing the algebra you could try to compute the slope by using a calculator. This exercise is about how you do that and what happens if you try (too hard).

Compute $\frac{\Delta y}{\Delta x}$ for various values of Δx :

$$\Delta x = 0.1, 0.01, 0.001, 10^{-6}, 10^{-12}.$$

As you choose Δx smaller your computed $\frac{\Delta y}{\Delta x}$ ought to get closer to the actual slope. Use at least 10 decimals and organize your results in a table like this:

| Δx | $f(a)$ | $f(a + \Delta x)$ | Δy | $\Delta y/\Delta x$ |
|------------|--------|-------------------|------------|---------------------|
| 0.1 | ... | ... | ... | ... |
| 0.01 | ... | ... | ... | ... |
| 0.001 | ... | ... | ... | ... |
| 10^{-6} | ... | ... | ... | ... |
| 10^{-12} | ... | ... | ... | ... |

Look carefully at the ratios $\Delta y/\Delta x$. Do they look like they are converging to some number? Compare the values of $\frac{\Delta y}{\Delta x}$ with the true value you got in the beginning of this problem.

30. Simplify the algebraic expressions you get when you compute Δy and $\Delta y/\Delta x$ for the following functions

(a) $y = x^2 - 2x + 1$

(b) $y = \frac{1}{x}$

(c) $y = 2^x$

31. Look ahead at Figure 3 in the next chapter. What is the derivative of $f(x) = x \cos \frac{\pi}{x}$ at the points A and B on the graph?

32. Suppose that some quantity y is a function of some other quantity x , and suppose that y is a mass, i.e. y is measured in pounds, and x is a length, measured in feet. What units do the increments Δy and Δx , and the derivative dy/dx have?

33. A tank is filling with water. The volume (in gallons) of water in the tank at time t (seconds) is $V(t)$. What units does the derivative $V'(t)$ have?

34. Group Problem.

Let $A(x)$ be the area of an equilateral triangle whose sides measure x inches.

(a) Show that $\frac{dA}{dx}$ has the units of a length.

(b) Which length does $\frac{dA}{dx}$ represent geometrically? [Hint: draw two equilateral triangles, one with side x and another with side $x + \Delta x$. Arrange the triangles so that they both have the origin as their lower left hand corner, and so their base is on the x -axis.]

35. Group Problem.

Let $A(x)$ be the area of a square with side x , and let $L(x)$ be the perimeter of the square (sum of the lengths of all its sides). Using the familiar formulas for $A(x)$ and $L(x)$ show that $A'(x) = \frac{1}{2}L(x)$.

Give a geometric interpretation that explains why $\Delta A \approx \frac{1}{2}L(x)\Delta x$ for small Δx .

36. Let $A(r)$ be the area enclosed by a circle of radius r , and let $L(r)$ be the length of the circle. Show that $A'(r) = L(r)$. (Use the familiar formulas from geometry for the area and perimeter of a circle.)

37. Let $V(r)$ be the volume enclosed by a sphere of radius r , and let $S(r)$ be its surface area. Show that $V'(r) = S(r)$. (Use the formulas $V(r) = \frac{4}{3}\pi r^3$ and $S(r) = 4\pi r^2$.)

Limits and Continuous Functions

1. Informal definition of limits

While it is easy to define precisely in a few words what a square root is (\sqrt{a} is the positive number whose square is a) the definition of the limit of a function runs over several terse lines, and most people don't find it very enlightening when they first see it. (See §2.) So we postpone this for a while and fine tune our intuition for another page.

1.1. Definition of limit (1st attempt). If f is some function then

$$\lim_{x \rightarrow a} f(x) = L$$

is read “the limit of $f(x)$ as x approaches a is L .” It means that if you choose values of x which are close **but not equal** to a , then $f(x)$ will be close to the value L ; moreover, $f(x)$ gets closer and closer to L as x gets closer and closer to a .

The following alternative notation is sometimes used

$$f(x) \rightarrow L \quad \text{as} \quad x \rightarrow a;$$

(read “ $f(x)$ approaches L as x approaches a ” or “ $f(x)$ goes to L as x goes to a ”.)

1.2. Example. If $f(x) = x + 3$ then

$$\lim_{x \rightarrow 4} f(x) = 7,$$

is true, because if you substitute numbers x close to 4 in $f(x) = x + 3$ the result will be close to 7.

1.3. Example: substituting numbers to guess a limit. What (if anything) is

$$\lim_{x \rightarrow 2} \frac{x^2 - 2x}{x^2 - 4}?$$

Here $f(x) = (x^2 - 2x)/(x^2 - 4)$ and $a = 2$.

We first try to substitute $x = 2$, but this leads to

$$f(2) = \frac{2^2 - 2 \cdot 2}{2^2 - 4} = \frac{0}{0}$$

which does not exist. Next we try to substitute values of x close but not equal to 2. Table 1 suggests that $f(x)$ approaches 0.5.

| x | $f(x)$ | x | $g(x)$ |
|----------|----------|----------|----------|
| 3.000000 | 0.600000 | 1.000000 | 1.009990 |
| 2.500000 | 0.555556 | 0.500000 | 1.009980 |
| 2.100000 | 0.512195 | 0.100000 | 1.009899 |
| 2.010000 | 0.501247 | 0.010000 | 1.008991 |
| 2.001000 | 0.500125 | 0.001000 | 1.000000 |

Table 1. Finding limits by substituting values of x “close to a .” (Values of $f(x)$ and $g(x)$ rounded to six decimals.)

1.4. Example: Substituting numbers can suggest the wrong answer. The previous example shows that our first definition of “limit” is not very precise, because it says “ x close to a ,” but how close is close enough? Suppose we had taken the function

$$g(x) = \frac{101\,000x}{100\,000x + 1}$$

and we had asked for the limit $\lim_{x \rightarrow 0} g(x)$.

Then substitution of some “small values of x ” could lead us to believe that the limit is $1.000\dots$. Only when you substitute even smaller values do you find that the limit is 0 (zero)!

See also problem [29](#).

2. The formal, authoritative, definition of limit

The informal description of the limit uses phrases like “closer and closer” and “really very small.” In the end we don’t really know what they mean, although they are suggestive. “Fortunately” there is a good definition, i.e. one which is unambiguous and can be used to settle any dispute about the question of whether $\lim_{x \rightarrow a} f(x)$ equals some number L or not. Here is the definition. It takes a while to digest, so read it once, look at the examples, do a few exercises, read the definition again. Go on to the next sections. Throughout the semester come back to this section and read it again.

2.1. Definition of $\lim_{x \rightarrow a} f(x) = L$. We say that L is the limit of $f(x)$ as $x \rightarrow a$, if

- (1) $f(x)$ need not be defined at $x = a$, but it must be defined for all other x in some interval which contains a .
- (2) for every $\varepsilon > 0$ one can find a $\delta > 0$ such that for all x in the domain of f one has

$$(8) \quad |x - a| < \delta \text{ implies } |f(x) - L| < \varepsilon.$$

Why the absolute values? The quantity $|x - y|$ is the distance between the points x and y on the number line, and one can measure how close x is to y by calculating $|x - y|$. The inequality $|x - y| < \delta$ says that “the distance between x and y is less than δ ,” or that “ x and y are closer than δ .”

What are ε and δ ? The quantity ε is how close you would like $f(x)$ to be to its limit L ; the quantity δ is how close you have to choose x to a to achieve this. To prove that $\lim_{x \rightarrow a} f(x) = L$ you must assume that someone has given you an unknown $\varepsilon > 0$, and then find a positive δ for which (8) holds. The δ you find will depend on ε .

2.2. Show that $\lim_{x \rightarrow 5} 2x + 1 = 11$. We have $f(x) = 2x + 1$, $a = 5$ and $L = 11$, and the question we must answer is “how close should x be to 5 if want to be sure that $f(x) = 2x + 1$ differs less than ε from $L = 11$?”

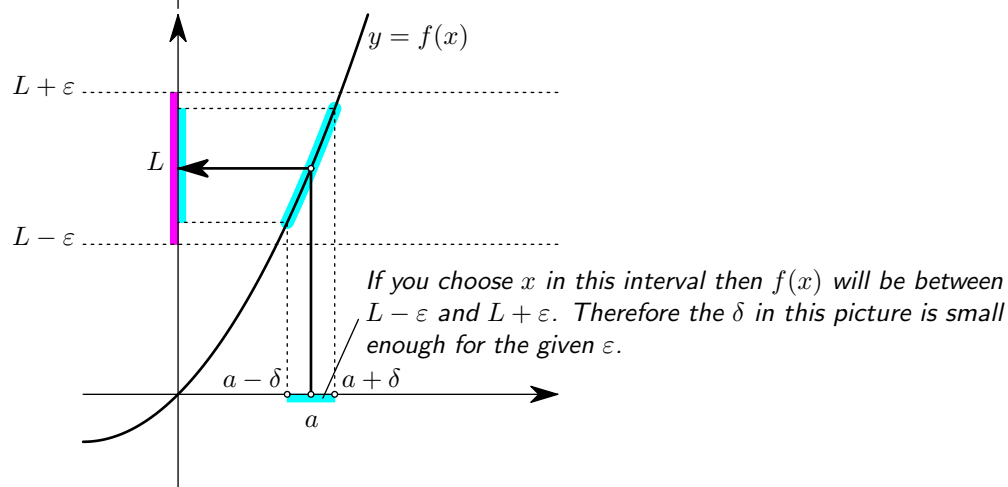
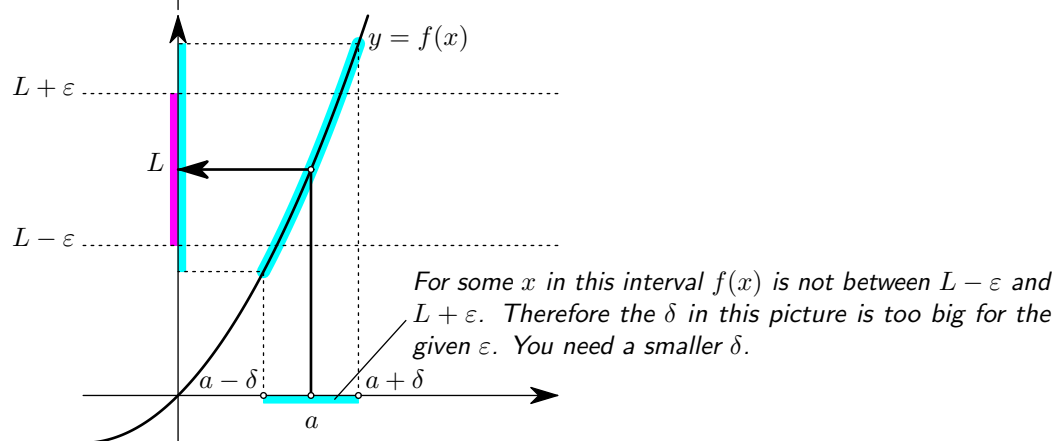
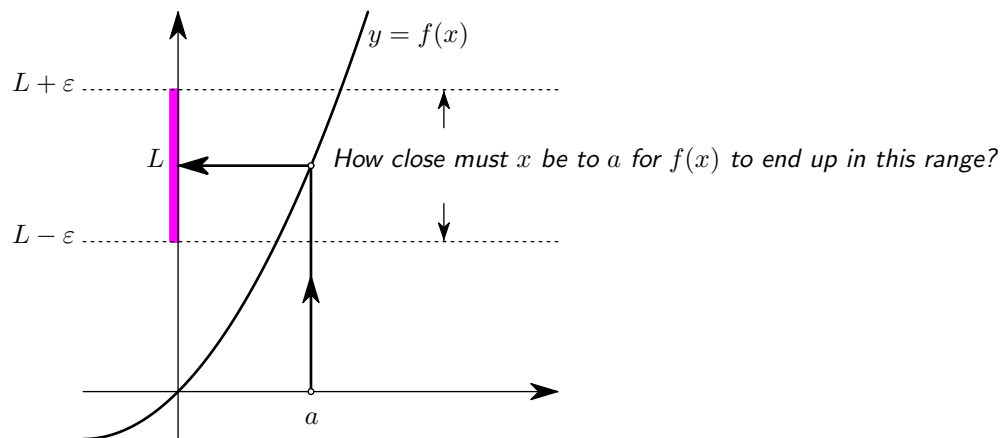
To figure this out we try to get an idea of how big $|f(x) - L|$ is:

$$|f(x) - L| = |(2x + 1) - 11| = |2x - 10| = 2 \cdot |x - 5| = 2 \cdot |x - a|.$$

So, if $2|x - a| < \varepsilon$ then we have $|f(x) - L| < \varepsilon$, i.e.

$$\text{if } |x - a| < \frac{1}{2}\varepsilon \text{ then } |f(x) - L| < \varepsilon.$$

We can therefore choose $\delta = \frac{1}{2}\varepsilon$. No matter what $\varepsilon > 0$ we are given our δ will also be positive, and if $|x - 5| < \delta$ then we can guarantee $|(2x + 1) - 11| < \varepsilon$. That shows that $\lim_{x \rightarrow 5} 2x + 1 = 11$.



2.3. The limit $\lim_{x \rightarrow 1} x^2 = 1$ and the “don’t choose $\delta > 1$ ” trick. We have $f(x) = x^2$, $a = 1$, $L = 1$, and again the question is, “how small should $|x - 1|$ be to guarantee $|x^2 - 1| < \varepsilon$?”

We begin by estimating the difference $|x^2 - 1|$

$$|x^2 - 1| = |(x - 1)(x + 1)| = |x + 1| \cdot |x - 1|.$$

Propagation of errors – another interpretation of ε and δ

According to the limit definition “ $\lim_{x \rightarrow R} \pi x^2 = A$ ” is true if *for every $\varepsilon > 0$ you can find a $\delta > 0$ such that $|x - R| < \delta$ implies $|\pi x^2 - A| < \varepsilon$* . Here's a more concrete situation in which ε and δ appear in exactly the same roles:

Suppose you are given a circle drawn on a piece of paper, and you want to know its area. You decide to measure its radius, R , and then compute the area of the circle by calculating

$$\text{Area} = \pi R^2.$$

The area is a function of the radius, and we'll call that function f :

$$f(x) = \pi x^2.$$

When you measure the radius R you will make an error, simply because you can never measure anything with infinite precision. Suppose that R is the real value of the radius, and that x is the number you measured. Then the size of the error you made is

$$\text{error in radius measurement} = |x - R|.$$

When you compute the area you also won't get the exact value: you would get $f(x) = \pi x^2$ instead of $A = f(R) = \pi R^2$. The error in your computed value of the area is

$$\text{error in area} = |f(x) - f(R)| = |f(x) - A|.$$

Now you can ask the following question:

Suppose you want to know the area with an error of at most ε , then what is the largest error that you can afford to make when you measure the radius?

The answer will be something like this: if you want the computed area to have an error of at most $|f(x) - A| < \varepsilon$, then the error in your radius measurement should satisfy $|x - R| < \delta$. You have to do the algebra with inequalities to compute δ when you know ε , as in the examples in this section.

You would expect that if your measured radius x is close enough to the real value R , then your computed area $f(x) = \pi x^2$ will be close to the real area A .

In terms of ε and δ this means that you would expect that no matter how accurately you want to know the area (i.e. how small you make ε) you can always achieve that precision by making the error in your radius measurement small enough (i.e. by making δ sufficiently small).

As x approaches 1 the factor $|x - 1|$ becomes small, and if the other factor $|x + 1|$ were a constant (e.g. 2 as in the previous example) then we could find δ as before, by dividing ε by that constant.

Here is a trick that allows you to replace the factor $|x + 1|$ with a constant. We hereby agree *that we always choose our δ so that $\delta \leq 1$* . If we do that, then we will always have

$$|x - 1| < \delta \leq 1, \text{ i.e. } |x - 1| < 1,$$

and x will always be between 0 and 2. Therefore

$$|x^2 - 1| = |x + 1| \cdot |x - 1| < 3|x - 1|.$$

If we now want to be sure that $|x^2 - 1| < \varepsilon$, then this calculation shows that we should require $3|x - 1| < \varepsilon$, i.e. $|x - 1| < \frac{1}{3}\varepsilon$. So we should choose $\delta \leq \frac{1}{3}\varepsilon$. We must also live up to our promise never to choose $\delta > 1$, so if we are handed an ε for which $\frac{1}{3}\varepsilon > 1$, then we choose $\delta = 1$ instead of $\delta = \frac{1}{3}\varepsilon$. To summarize, we are going to choose

$$\delta = \text{the smaller of } 1 \text{ and } \frac{1}{3}\varepsilon.$$

We have shown that if you choose δ this way, then $|x - 1| < \delta$ implies $|x^2 - 1| < \varepsilon$, no matter what $\varepsilon > 0$ is.

The expression “the smaller of a and b ” shows up often, and is abbreviated to $\min(a, b)$. We could therefore say that in this problem we will choose δ to be

$$\delta = \min(1, \frac{1}{3}\varepsilon).$$

2.4. Show that $\lim_{x \rightarrow 4} 1/x = 1/4$. Solution: We apply the definition with $a = 4$, $L = 1/4$ and $f(x) = 1/x$. Thus, for any $\varepsilon > 0$ we try to show that if $|x - 4|$ is small enough then one has $|f(x) - 1/4| < \varepsilon$.

We begin by estimating $|f(x) - 1/4|$ in terms of $|x - 4|$:

$$|f(x) - 1/4| = \left| \frac{1}{x} - \frac{1}{4} \right| = \left| \frac{4 - x}{4x} \right| = \frac{|x - 4|}{|4x|} = \frac{1}{|4x|} |x - 4|.$$

As before, things would be easier if $1/|4x|$ were a constant. To achieve that we again agree not to take $\delta > 1$. If we always have $\delta \leq 1$, then we will always have $|x - 4| < 1$, and hence $3 < x < 5$. How large can $1/|4x|$ be in this situation? Answer: the quantity $1/|4x|$ increases as you decrease x , so if $3 < x < 5$ then it will never be larger than $1/(4 \cdot 3) = \frac{1}{12}$.

We see that if we never choose $\delta > 1$, we will always have

$$|f(x) - \frac{1}{4}| \leq \frac{1}{12} |x - 4| \quad \text{for } |x - 4| < \delta.$$

To guarantee that $|f(x) - \frac{1}{4}| < \varepsilon$ we could therefore require

$$\frac{1}{12} |x - 4| < \varepsilon, \quad \text{i.e. } |x - 4| < 12\varepsilon.$$

Hence if we choose $\delta = 12\varepsilon$ or any smaller number, then $|x - 4| < \delta$ implies $|f(x) - 1/4| < \varepsilon$. Of course we have to honor our agreement never to choose $\delta > 1$, so our choice of δ is

$$\delta = \text{the smaller of } 1 \text{ and } 12\varepsilon = \min(1, 12\varepsilon).$$

3. Exercises

38. Group Problem.

Joe offers to make square sheets of paper for Bruce. Given $x > 0$ Joe plans to mark off a length x and cut out a square of side x . Bruce asks Joe for a square with area 4 square foot. Joe tells Bruce that he can't measure **exactly** 2 foot and the area of the square he produces will only be approximately 4 square foot. Bruce doesn't mind as long as the area of the square doesn't differ more than 0.01 square foot from what he really asked for (namely, 4 square foot).

(a) What is the biggest error Joe can afford to make when he marks off the length x ?

(b) Jen also wants square sheets, with area 4 square feet. However, she needs the error in the area to be less than 0.00001 square foot. (She's paying).

How accurate must Joe measure the side of the squares he's going to cut for Jen?

Use the ε - δ definition to prove the following limits

39. $\lim_{x \rightarrow 1} 2x - 4 = 6$

40. $\lim_{x \rightarrow 2} x^2 = 4$.

41. $\lim_{x \rightarrow 2} x^2 - 7x + 3 = -7$

42. $\lim_{x \rightarrow 3} x^3 = 27$

43. $\lim_{x \rightarrow 2} x^3 + 6x^2 = 32$.

44. $\lim_{x \rightarrow 4} \sqrt{x} = 2$.

45. $\lim_{x \rightarrow 3} \sqrt{x+6} = 9$.

46. $\lim_{x \rightarrow 2} \frac{1+x}{4+x} = \frac{1}{2}$.

47. $\lim_{x \rightarrow 1} \frac{2-x}{4-x} = \frac{1}{3}$.

48. $\lim_{x \rightarrow 3} \frac{x}{6-x} = 1$.

49. $\lim_{x \rightarrow 0} \sqrt{|x|} = 0$

50. Group Problem.

(Joe goes cubic.) Joe is offering to build cubes of side x . Airline regulations allow you take a cube on board provided its volume and surface area add up to less than 33 (everything measured in feet). For instance, a cube with 2 foot sides has volume+area equal to $2^3 + 6 \times 2^2 = 32$.

If you ask Joe to build a cube whose volume plus total surface area is 32 cubic feet with an error of at most ε , then what error can he afford to make when he measures the side of the cube he's making?

51. Our definition of a derivative in (7) contains a limit. What is the function " f " there, and what is the variable?

4. Variations on the limit theme

Not all limits are "for $x \rightarrow a$." here we describe some possible variations on the concept of limit.

4.1. Left and right limits. When we let “ x approach a ” we allow x to be both larger or smaller than a , as long as x gets close to a . If we explicitly want to study the behaviour of $f(x)$ as x approaches a through values larger than a , then we write

$$\lim_{x \searrow a} f(x) \text{ or } \lim_{x \rightarrow a+} f(x) \text{ or } \lim_{x \rightarrow a+0} f(x) \text{ or } \lim_{x \rightarrow a, x > a} f(x).$$

All four notations are in use. Similarly, to designate the value which $f(x)$ approaches as x approaches a through values below a one writes

$$\lim_{x \nearrow a} f(x) \text{ or } \lim_{x \rightarrow a-} f(x) \text{ or } \lim_{x \rightarrow a-0} f(x) \text{ or } \lim_{x \rightarrow a, x < a} f(x).$$

The precise definition of right limits goes like this:

4.2. Definition of right-limits. *Let f be a function. Then*

$$(9) \quad \lim_{x \searrow a} f(x) = L.$$

means that for every $\varepsilon > 0$ one can find a $\delta > 0$ such that

$$a < x < a + \delta \implies |f(x) - L| < \varepsilon$$

holds for all x in the domain of f .

The left-limit, i.e. the one-sided limit in which x approaches a through values less than a is defined in a similar way. The following theorem tells you how to use one-sided limits to decide if a function $f(x)$ has a limit at $x = a$.

4.3. Theorem. *If both one-sided limits*

$$\lim_{x \searrow a} f(x) = L_+, \text{ and } \lim_{x \nearrow a} f(x) = L_-$$

exist, then

$$\lim_{x \rightarrow a} f(x) \text{ exists } \iff L_+ = L_-.$$

In other words, if a function has both left- and right-limits at some $x = a$, then that function has a limit at $x = a$ if the left- and right-limits are equal.

4.4. Limits at infinity. Instead of letting x approach some finite number, one can let x become “larger and larger” and ask what happens to $f(x)$. If there is a number L such that $f(x)$ gets arbitrarily close to L if one chooses x sufficiently large, then we write

$$\lim_{x \rightarrow \infty} f(x) = L, \text{ or } \lim_{x \uparrow \infty} f(x) = L, \text{ or } \lim_{x \nearrow \infty} f(x) = L.$$

(“The limit for x going to infinity is L .”)

4.5. Example – Limit of $1/x$. The larger you choose x , the smaller its reciprocal $1/x$ becomes. Therefore, it seems reasonable to say

$$\lim_{x \rightarrow \infty} \frac{1}{x} = 0.$$

Here is the precise definition:

4.6. Definition of limit at ∞ . *Let f be some function which is defined on some interval $x_0 < x < \infty$. If there is a number L such that for every $\varepsilon > 0$ one can find an A such that*

$$x > A \implies |f(x) - L| < \varepsilon$$

for all x , then we say that the limit of $f(x)$ for $x \rightarrow \infty$ is L .

The definition is very similar to the original definition of the limit. Instead of δ which specifies how close x should be to a , we now have a number A which says how large x should be, which is a way of saying “how close x should be to infinity.”

4.7. Example – Limit of $1/x$ (again) . To *prove* that $\lim_{x \rightarrow \infty} 1/x = 0$ we apply the definition to $f(x) = 1/x$, $L = 0$.

For given $\varepsilon > 0$ we need to show that

$$(10) \quad \left| \frac{1}{x} - L \right| < \varepsilon \text{ for all } x > A$$

provided we choose the right A .

How do we choose A ? A is not allowed to depend on x , but it may depend on ε .

If we assume for now that we will only consider positive values of x , then (10) simplifies to

$$\frac{1}{x} < \varepsilon$$

which is equivalent to

$$x > \frac{1}{\varepsilon}.$$

This tells us how to choose A . Given any positive ε , we will simply choose

$$A = \frac{1}{\varepsilon}$$

Then one has $|\frac{1}{x} - 0| = \frac{1}{x} < \varepsilon$ for all $x > A$. Hence we have proved that $\lim_{x \rightarrow \infty} 1/x = 0$.

5. Properties of the Limit

The precise definition of the limit is not easy to use, and fortunately we won't use it very often in this class. Instead, there are a number of properties that limits have which allow you to compute them without having to resort to "epsilon-ness."

The following properties also apply to the variations on the limit from 4. I.e. the following statements remain true if one replaces each limit by a one-sided limit, or a limit for $x \rightarrow \infty$.

Limits of constants and of x . If a and c are constants, then

$$(P_1) \quad \lim_{x \rightarrow a} c = c$$

and

$$(P_2) \quad \lim_{x \rightarrow a} x = a.$$

Limits of sums, products and quotients. Let F_1 and F_2 be two given functions whose limits for $x \rightarrow a$ we know,

$$\lim_{x \rightarrow a} F_1(x) = L_1, \quad \lim_{x \rightarrow a} F_2(x) = L_2.$$

Then

$$(P_3) \quad \lim_{x \rightarrow a} (F_1(x) + F_2(x)) = L_1 + L_2,$$

$$(P_4) \quad \lim_{x \rightarrow a} (F_1(x) - F_2(x)) = L_1 - L_2,$$

$$(P_5) \quad \lim_{x \rightarrow a} (F_1(x) \cdot F_2(x)) = L_1 \cdot L_2$$

Finally, if $\lim_{x \rightarrow a} F_2(x) \neq 0$,

$$(P_6) \quad \lim_{x \rightarrow a} \frac{F_1(x)}{F_2(x)} = \frac{L_1}{L_2}.$$

In other words the limit of the sum is the sum of the limits, etc. One can prove these laws using the definition of limit in §2 but we will not do this here. However, I hope these laws seem like common sense: if, for x close to a , the quantity $F_1(x)$ is close to L_1 and $F_2(x)$ is close to L_2 , then certainly $F_1(x) + F_2(x)$ should be close to $L_1 + L_2$.

There are two more properties of limits which we will add to this list later on. They are the “Sandwich Theorem” (§9) and the substitution theorem (§10).

6. Examples of limit computations

6.1. Find $\lim_{x \rightarrow 2} x^2$. One has

$$\begin{aligned}\lim_{x \rightarrow 2} x^2 &= \lim_{x \rightarrow 2} x \cdot x \\ &= \left(\lim_{x \rightarrow 2} x \right) \cdot \left(\lim_{x \rightarrow 2} x \right) && \text{by } (P_5) \\ &= 2 \cdot 2 = 4.\end{aligned}$$

Similarly,

$$\begin{aligned}\lim_{x \rightarrow 2} x^3 &= \lim_{x \rightarrow 2} x \cdot x^2 \\ &= \left(\lim_{x \rightarrow 2} x \right) \cdot \left(\lim_{x \rightarrow 2} x^2 \right) && (P_5) \text{ again} \\ &= 2 \cdot 4 = 8,\end{aligned}$$

and, by (P_4)

$$\lim_{x \rightarrow 2} x^2 - 1 = \lim_{x \rightarrow 2} x^2 - \lim_{x \rightarrow 2} 1 = 4 - 1 = 3,$$

and, by (P_4) again,

$$\lim_{x \rightarrow 2} x^3 - 1 = \lim_{x \rightarrow 2} x^3 - \lim_{x \rightarrow 2} 1 = 8 - 1 = 7,$$

Putting all this together, one gets

$$\lim_{x \rightarrow 2} \frac{x^3 - 1}{x^2 - 1} = \frac{2^3 - 1}{2^2 - 1} = \frac{8 - 1}{4 - 1} = \frac{7}{3}$$

because of (P_6) . To apply (P_6) we must check that the denominator (“ L_2 ”) is not zero. Since the denominator is 3 everything is OK, and we were allowed to use (P_6) .

6.2. Try the examples 1.3 and 1.4 using the limit properties. To compute $\lim_{x \rightarrow 2} (x^2 - 2x)/(x^2 - 4)$ we first use the limit properties to find

$$\lim_{x \rightarrow 2} x^2 - 2x = 0 \text{ and } \lim_{x \rightarrow 2} x^2 - 4 = 0.$$

to complete the computation we would like to apply the last property (P_6) about quotients, but this would give us

$$\lim_{x \rightarrow 2} f(x) = \frac{0}{0}.$$

The denominator is zero, so we were not allowed to use (P_6) (and the result doesn’t mean anything anyway). We have to do something else.

The function we are dealing with is a **rational function**, which means that it is the quotient of two polynomials. For such functions there is an algebra trick which always allows you to compute the limit even if you first get $\frac{0}{0}$. The thing to do is to divide numerator and denominator by $x - 2$. In our case we have

$$x^2 - 2x = (x - 2) \cdot x, \quad x^2 - 4 = (x - 2) \cdot (x + 2)$$

so that

$$\lim_{x \rightarrow 2} f(x) = \lim_{x \rightarrow 2} \frac{(x - 2) \cdot x}{(x - 2) \cdot (x + 2)} = \lim_{x \rightarrow 2} \frac{x}{x + 2}.$$

After this simplification we **can** use the properties (P_{\dots}) to compute

$$\lim_{x \rightarrow 2} f(x) = \frac{2}{2 + 2} = \frac{1}{2}.$$

6.3. Example – Find $\lim_{x \rightarrow 2} \sqrt{x}$. Of course, you would think that $\lim_{x \rightarrow 2} \sqrt{x} = \sqrt{2}$ and you can indeed prove this using δ & ε (See problem 44.) But is there an easier way? There is nothing in the limit properties which tells us how to deal with a square root, and using them we can't even prove that there is a limit. However, if you **assume** that the limit exists then the limit properties allow us to find this limit.

The argument goes like this: suppose that there is a number L with

$$\lim_{x \rightarrow 2} \sqrt{x} = L.$$

Then property (P_5) implies that

$$L^2 = \left(\lim_{x \rightarrow 2} \sqrt{x}\right) \cdot \left(\lim_{x \rightarrow 2} \sqrt{x}\right) = \lim_{x \rightarrow 2} \sqrt{x} \cdot \sqrt{x} = \lim_{x \rightarrow 2} x = 2.$$

In other words, $L^2 = 2$, and hence L must be either $\sqrt{2}$ or $-\sqrt{2}$. We can reject the latter because whatever x does, its squareroot is always a positive number, and hence it can never “get close to” a negative number like $-\sqrt{2}$.

Our conclusion: if the limit exists, then

$$\lim_{x \rightarrow 2} \sqrt{x} = \sqrt{2}.$$

The result is not surprising: if x gets close to 2 then \sqrt{x} gets close to $\sqrt{2}$.

6.4. Example – The derivative of \sqrt{x} at $x = 2$. Find

$$\lim_{x \rightarrow 2} \frac{\sqrt{x} - \sqrt{2}}{x - 2}$$

assuming the result from the previous example.

Solution: The function is a fraction whose numerator and denominator vanish when $x = 2$, i.e. the limit is of the form $\frac{0}{0}$. We use the same algebra trick as before, namely we factor numerator and denominator:

$$\frac{\sqrt{x} - \sqrt{2}}{x - 2} = \frac{\sqrt{x} - \sqrt{2}}{(\sqrt{x} - \sqrt{2})(\sqrt{x} + \sqrt{2})} = \frac{1}{\sqrt{x} + \sqrt{2}}.$$

Now one can use the limit properties to compute

$$\lim_{x \rightarrow 2} \frac{\sqrt{x} - \sqrt{2}}{x - 2} = \lim_{x \rightarrow 2} \frac{1}{\sqrt{x} + \sqrt{2}} = \frac{1}{2\sqrt{2}} = \frac{\sqrt{2}}{4}.$$

6.5. Limit as $x \rightarrow \infty$ of rational functions. A rational function is the quotient of two polynomials, so

$$(11) \quad R(x) = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}.$$

We have seen that

$$\lim_{x \rightarrow \infty} \frac{1}{x} = 0$$

We even proved this in example 4.7. Using this you can find the limit at ∞ for any rational function $R(x)$ as in (11). One could turn the outcome of the calculation of $\lim_{x \rightarrow \infty} R(x)$ into a recipe/formula involving the degrees n and m of the numerator and denominator, and also their coefficients a_i , b_j , which students would then memorize, but it is better to remember “the trick.”

To find $\lim_{x \rightarrow \infty} R(x)$ divide numerator and denominator by x^m (the highest power of x occurring in the denominator).

For example, let's compute

$$\lim_{x \rightarrow \infty} \frac{3x^2 + 3}{5x^2 + 7x - 39}.$$

Remember the trick and divide top and bottom by x^2 , and you get

$$\begin{aligned}\lim_{x \rightarrow \infty} \frac{3x^2 + 3}{5x^2 + 7x - 39} &= \lim_{x \rightarrow \infty} \frac{3 + 3/x^2}{5 + 7/x - 39/x^2} \\ &= \frac{\lim_{x \rightarrow \infty} 3 + 3/x^2}{\lim_{x \rightarrow \infty} 5 + 7/x - 39/x^2} \\ &= \frac{3}{5}\end{aligned}$$

Here we have used the limit properties (P_*) to break the limit down into little pieces like $\lim_{x \rightarrow \infty} 39/x^2$ which we can compute as follows

$$\lim_{x \rightarrow \infty} 39/x^2 = \lim_{x \rightarrow \infty} 39 \cdot \left(\frac{1}{x}\right)^2 = \left(\lim_{x \rightarrow \infty} 39\right) \cdot \left(\lim_{x \rightarrow \infty} \frac{1}{x}\right)^2 = 39 \cdot 0^2 = 0.$$

6.6. Another example with a rational function . Compute

$$\lim_{x \rightarrow \infty} \frac{x}{x^3 + 5}.$$

We apply “the trick” again and divide numerator and denominator by x^3 . This leads to

$$\lim_{x \rightarrow \infty} \frac{x}{x^3 + 5} = \lim_{x \rightarrow \infty} \frac{1/x^2}{1 + 5/x^3} = \frac{\lim_{x \rightarrow \infty} 1/x^2}{\lim_{x \rightarrow \infty} 1 + 5/x^3} = \frac{0}{1} = 0.$$

To show all possible ways a limit of a rational function can turn out we should do yet another example, but that one belongs in the next section (see example 7.6.)

7. When limits fail to exist

In the last couple of examples we worried about the possibility that a limit $\lim_{x \rightarrow a} g(x)$ actually might not exist. This can actually happen, and in this section we’ll see a few examples of what failed limits look like. First let’s agree on what we will call a “failed limit.”

7.1. Definition. *If there is no number L such that $\lim_{x \rightarrow a} f(x) = L$, then we say that the limit $\lim_{x \rightarrow a} f(x)$ does not exist.*

7.2. The sign function near $x = 0$. The “sign function¹” is defined by

$$\text{sign}(x) = \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ 1 & \text{for } x > 0 \end{cases}$$

Note that “the sign of zero” is defined to be zero. But does the sign function have a limit at $x = 0$, i.e. does $\lim_{x \rightarrow 0} \text{sign}(x)$ exist? And is it also zero? The answers are **no** and **no**, and here is why: suppose that for some number L one had

$$\lim_{x \rightarrow 0} \text{sign}(x) = L,$$

then since for arbitrary small positive values of x one has $\text{sign}(x) = +1$ one would think that $L = +1$. But for arbitrarily small negative values of x one has $\text{sign}(x) = -1$, so one would conclude that $L = -1$. But one number L can’t be both $+1$ and -1 at the same time, so there is no such L , i.e. there is no limit.

$$\lim_{x \rightarrow 0} \text{sign}(x) \text{ does not exist.}$$

¹Some people don’t like the notation $\text{sign}(x)$, and prefer to write

$$g(x) = \frac{x}{|x|}$$

instead of $g(x) = \text{sign}(x)$. If you think about this formula for a moment you’ll see that $\text{sign}(x) = x/|x|$ for all $x \neq 0$. When $x = 0$ the quotient $x/|x|$ is of course not defined.

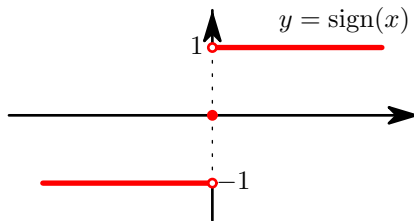


Figure 1. The sign function.

In this example the one-sided limits do exist, namely,

$$\lim_{x \searrow 0} \text{sign}(x) = 1 \text{ and } \lim_{x \nearrow 0} \text{sign}(x) = -1.$$

All this says is that when x approaches 0 through positive values, its sign approaches +1, while if x goes to 0 through negative values, then its sign approaches -1.

7.3. The example of the backward sine. Contemplate the limit as $x \rightarrow 0$ of the “backward sine,” i.e.

$$\lim_{x \rightarrow 0} \sin\left(\frac{\pi}{x}\right).$$

When $x = 0$ the function $f(x) = \sin(\pi/x)$ is not defined, because its definition involves division by x . What happens to $f(x)$ as $x \rightarrow 0$? First, π/x becomes larger and larger (“goes to infinity”) as $x \rightarrow 0$. Then, taking the sine, we see that $\sin(\pi/x)$ oscillates between +1 and -1 infinitely often as $x \rightarrow 0$. This means that $f(x)$ gets close to any number between -1 and +1 as $x \rightarrow 0$, but that the function $f(x)$ *never stays close* to any particular value because it keeps oscillating up and down.

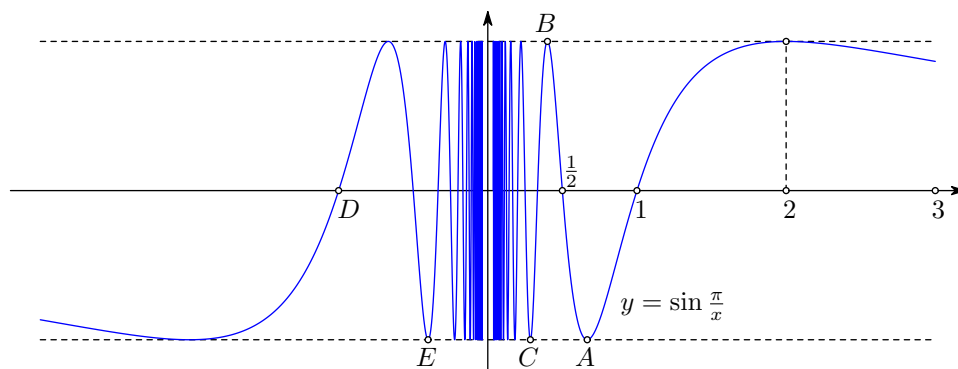


Figure 2. Graph of $y = \sin \frac{\pi}{x}$ for $-3 < x < 3$, $x \neq 0$.

Here again, the limit $\lim_{x \rightarrow 0} f(x)$ does not exist. We have arrived at this conclusion by only considering what $f(x)$ does for small positive values of x . So the limit fails to exist in a stronger way than in the example of the sign-function. There, even though the limit didn’t exist, the one-sided limits existed. In the present example we see that even the one-sided limit

$$\lim_{x \searrow 0} \sin \frac{\pi}{x}$$

does not exist.

7.4. Trying to divide by zero using a limit. The expression $1/0$ is not defined, but what about

$$\lim_{x \rightarrow 0} \frac{1}{x}?$$

This limit also does not exist. Here are two reasons:

It is common wisdom that if you divide by a small number you get a large number, so as $x \searrow 0$ the quotient $1/x$ will not be able to stay close to any particular finite number, and the limit can't exist.

“Common wisdom” is not always a reliable tool in mathematical proofs, so here is a better argument. The limit can't exist, because that would contradict the limit properties $(P_1) \cdots (P_6)$. Namely, suppose that there were an number L such that

$$\lim_{x \rightarrow 0} \frac{1}{x} = L.$$

Then the limit property (P_5) would imply that

$$\lim_{x \rightarrow 0} \left(\frac{1}{x} \cdot x \right) = \left(\lim_{x \rightarrow 0} \frac{1}{x} \right) \cdot \left(\lim_{x \rightarrow 0} x \right) = L \cdot 0 = 0.$$

On the other hand $\frac{1}{x} \cdot x = 1$ so the above limit should be 1! A number can't be both 0 and 1 at the same time, so we have a contradiction. The assumption that $\lim_{x \rightarrow 0} 1/x$ exists is to blame, so it must go.

7.5. Using limit properties to show a limit does *not* exist. The limit properties tell us how to prove that certain limits exist (and how to compute them). Although it is perhaps not so obvious at first sight, they also allow you to prove that certain limits do not exist. The previous example shows one instance of such use. Here is another.

Property (P_3) says that if both $\lim_{x \rightarrow a} g(x)$ and $\lim_{x \rightarrow a} h(x)$ exist then $\lim_{x \rightarrow a} g(x) + h(x)$ also must exist. You can turn this around and say that if $\lim_{x \rightarrow a} g(x) + h(x)$ does not exist then either $\lim_{x \rightarrow a} g(x)$ or $\lim_{x \rightarrow a} h(x)$ does not exist (or both limits fail to exist).

For instance, the limit

$$\lim_{x \rightarrow 0} \frac{1}{x} - x$$

can't exist, for if it did, then the limit

$$\lim_{x \rightarrow 0} \frac{1}{x} = \lim_{x \rightarrow 0} \left(\frac{1}{x} - x + x \right) = \lim_{x \rightarrow 0} \left(\frac{1}{x} - x \right) + \lim_{x \rightarrow 0} x$$

would also have to exist, and we know $\lim_{x \rightarrow 0} \frac{1}{x}$ doesn't exist.

7.6. Limits at ∞ which don't exist. If you let x go to ∞ , then x will not get “closer and closer” to any particular number L , so it seems reasonable to guess that

$$\lim_{x \rightarrow \infty} x \text{ does not exist.}$$

One can prove this from the limit definition (and see exercise 72).

Let's consider

$$L = \lim_{x \rightarrow \infty} \frac{x^2 + 2x - 1}{x + 2}.$$

Once again we divide numerator and denominator by the highest power in the denominator (i.e. x)

$$L = \lim_{x \rightarrow \infty} \frac{x + 2 - \frac{1}{x}}{1 + 2/x}$$

Here the denominator has a limit ('tis 1), but the numerator does not, for if $\lim_{x \rightarrow \infty} x + 2 - \frac{1}{x}$ existed then, since $\lim_{x \rightarrow \infty} (2 - 1/x) = 2$ exists,

$$\lim_{x \rightarrow \infty} x = \lim_{x \rightarrow \infty} \left[\left(x + 2 - \frac{1}{x} \right) - \left(2 - \frac{1}{x} \right) \right]$$

would also have to exist, and $\lim_{x \rightarrow \infty} x$ doesn't exist.

So we see that L is the limit of a fraction in which the denominator has a limit, but the numerator does not. In this situation the limit L itself can never exist. If it did, then

$$\lim_{x \rightarrow \infty} \left(x + 2 - \frac{1}{x} \right) = \lim_{x \rightarrow \infty} \frac{x + 2 - \frac{1}{x}}{1 + 2/x} \cdot (1 + 2/x)$$

would also have to have a limit.

8. What's in a name?

There is a big difference between the variables x and a in the formula

$$\lim_{x \rightarrow a} 2x + 1,$$

namely a is a **free variable**, while x is a **dummy variable** (or “placeholder” or a “bound variable.”)

The difference between these two kinds of variables is this:

- if you replace a dummy variable in some formula consistently by some other variable then the value of the formula does not change. On the other hand, it never makes sense to substitute a number for a dummy variable.
- the value of the formula may depend on the value of the free variable.

To understand what this means consider the example $\lim_{x \rightarrow a} 2x + 1$ again. The limit is easy to compute:

$$\lim_{x \rightarrow a} 2x + 1 = 2a + 1.$$

If we replace x by, say u (systematically) then we get

$$\lim_{u \rightarrow a} 2u + 1$$

which is again equal to $2a + 1$. This computation says that *if some number gets close to a then two times that number plus one gets close to $2a + 1$* . This is a very wordy way of expressing the formula, and you can shorten things by giving a name (like x or u) to the number which approaches a . But the result of our computation shouldn't depend on the name we choose, i.e. it doesn't matter if we call it x or u .

Since the name of the variable x doesn't matter it is called a dummy variable. Some prefer to call x a bound variable, meaning that in

$$\lim_{x \rightarrow a} 2x + 1$$

the x in the expression $2x + 1$ is bound to the x written underneath the limit – you can't change one without changing the other.

Substituting a number for a dummy variable usually leads to complete nonsense. For instance, let's try setting $x = 3$ in our limit, i.e. what is

$$\lim_{3 \rightarrow a} 2 \cdot 3 + 1?$$

Of course $2 \cdot 3 + 1 = 7$, but what does 7 do when 3 gets closer and closer to the number a ? That's a silly question, because 3 is a constant and it doesn't “get closer” to some other number like a ! If you ever see 3 get closer to another number then it's time to take a vacation.

On the other hand the variable a is free: you can assign it particular values, and its value will affect the value of the limit. For instance, if we set $a = 3$ (but leave x alone) then we get

$$\lim_{x \rightarrow 3} 2x + 1$$

and there's nothing strange about that (the limit is $2 \cdot 3 + 1 = 7$, no problem.) You could substitute other values of a and you would get a different answer. In general you get $2a + 1$.

9. Limits and Inequalities

This section has two theorems which let you compare limits of different functions. The properties in these theorems are not formulas that allow you to compute limits like the properties $(P_1) \dots (P_6)$ from §5. Instead, they allow you to **reason** about limits, i.e. they let you say that this or that limit is positive, or that it must be the same as some other limit which you find easier to think about.

The first theorem should not surprise you – all it says is that bigger functions have bigger limits.

9.1. Theorem. *Let f and g be functions whose limits for $x \rightarrow a$ exist, and assume that $f(x) \leq g(x)$ holds for all x . Then*

$$\lim_{x \rightarrow a} f(x) \leq \lim_{x \rightarrow a} g(x).$$

A useful special case arises when you set $f(x) = 0$. The theorem then says that if a function g never has negative values, then its limit will also never be negative.

The statement may seem obvious, but it still needs a proof, starting from the ε - δ definition of limit. This will be done in lecture.

Here is the second theorem about limits and inequalities.

9.2. The Sandwich Theorem. *Suppose that*

$$f(x) \leq g(x) \leq h(x)$$

(for all x) and that

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x).$$

Then

$$\lim_{x \rightarrow a} g(x) = \lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x).$$

The theorem is useful when you want to know the limit of g , and when you can **sandwich** it between two functions f and h whose limits are easier to compute. The Sandwich Theorem looks like the first theorem of this section, but there is an important difference: in the Sandwich Theorem you don't have to assume that the limit of g exists. The inequalities $f \leq g \leq h$ combined with the circumstance that f and h have the same limit are enough to guarantee that the limit of g exists.

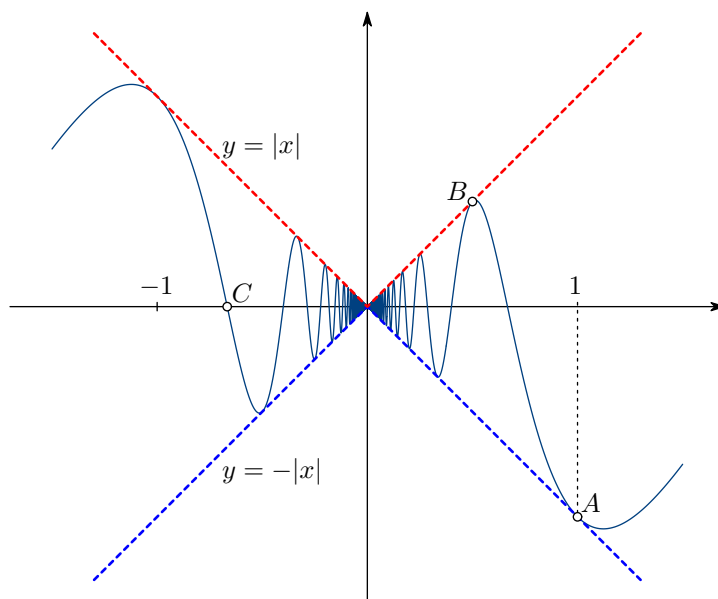


Figure 3. Graphs of $|x|$, $-|x|$ and $x \cos \frac{\pi}{x}$ for $-1.2 < x < 1.2$

9.3. Example: a Backward Cosine Sandwich. The Sandwich Theorem says that if the function $g(x)$ is sandwiched between two functions $f(x)$ and $h(x)$ and the limits of the outside functions f and h exist and are equal, then the limit of the inside function g exists and equals this common value. For example

$$-|x| \leq x \cos \frac{1}{x} \leq |x|$$

since the cosine is always between -1 and 1 . Since

$$\lim_{x \rightarrow 0} -|x| = \lim_{x \rightarrow 0} |x| = 0$$

the sandwich theorem tells us that

$$\lim_{x \rightarrow 0} x \cos \frac{1}{x} = 0.$$

Note that the limit $\lim_{x \rightarrow 0} \cos(1/x)$ does **not** exist, for the same reason that the “backward sine” did not have a limit for $x \rightarrow 0$ (see example 7.3). Multiplying with x changed that.

10. Continuity

10.1. Definition. A function g is **continuous** at a if

$$(12) \quad \lim_{x \rightarrow a} g(x) = g(a)$$

A function is continuous if it is continuous at every a in its domain.

Note that when we say that a function is continuous on some interval it is understood that the domain of the function includes that interval. For example, the function $f(x) = 1/x^2$ is continuous on the interval $1 < x < 5$ but is **not** continuous on the interval $-1 < x < 1$.

10.2. Polynomials are continuous. For instance, let us show that $P(x) = x^2 + 3x$ is continuous at $x = 2$. To show that you have to prove that

$$\lim_{x \rightarrow 2} P(x) = P(2),$$

i.e.

$$\lim_{x \rightarrow 2} x^2 + 3x = 2^2 + 3 \cdot 2.$$

You can do this two ways: using the definition with ε and δ (i.e. the hard way), or using the limit properties $(P_1) \dots (P_6)$ from §5 (just as good, and easier, even though it still takes a few lines to write it out – do both!)

10.3. Rational functions are continuous. Let $R(x) = \frac{P(x)}{Q(x)}$ be a rational function, and let a be any number in the domain of R , i.e. any number for which $Q(a) \neq 0$. Then one has

$$\begin{aligned} \lim_{x \rightarrow a} R(x) &= \lim_{x \rightarrow a} \frac{P(x)}{Q(x)} \\ &= \frac{\lim_{x \rightarrow a} P(x)}{\lim_{x \rightarrow a} Q(x)} && \text{property } (P_6) \\ &= \frac{P(a)}{Q(a)} && P \text{ and } Q \text{ are continuous} \\ &= R(a). \end{aligned}$$

This shows that R is indeed continuous at a .

10.4. Some discontinuous functions. If $\lim_{x \rightarrow a} g(x)$ does not exist, then it certainly cannot be equal to $g(a)$, and therefore any failed limit provides an example of a discontinuous function.

For instance, the sign function $g(x) = \text{sign}(x)$ from example ?? is not continuous at $x = 0$.

Is the backward sine function $g(x) = \sin(1/x)$ from example ?? also discontinuous at $x = 0$? No, it is not, for two reasons: first, the limit $\lim_{x \rightarrow 0} \sin(1/x)$ does not exist, and second, we haven’t even defined the function $g(x)$ at $x = 0$, so even if the limit existed, we would have no value $g(0)$ to compare it with.

10.5. How to make functions discontinuous. Here is a discontinuous function:

$$f(x) = \begin{cases} x^2 & \text{if } x \neq 3, \\ 47 & \text{if } x = 3. \end{cases}$$

In other words, we take a continuous function like $g(x) = x^2$, and change its value somewhere, e.g. at $x = 3$. Then

$$\lim_{x \rightarrow 3} f(x) = 9 \neq 47 = f(3).$$

The reason that the limit is 9 is that our new function $f(x)$ coincides with our old continuous function $g(x)$ for all x except $x = 3$. Therefore the limit of $f(x)$ as $x \rightarrow 3$ is the same as the limit of $g(x)$ as $x \rightarrow 3$, and since g is continuous this is $g(3) = 9$.

10.6. Sandwich in a bow tie. We return to the function from example ???. Consider

$$f(x) = \begin{cases} x \cos\left(\frac{1}{x}\right) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0 \end{cases}$$

Then f is continuous at $x = 0$ by the Sandwich Theorem (see Example ??).

If we change the definition of f by picking a different value at $x = 0$ the new function will not be continuous, since changing f at $x = 0$ does not change the limit $\lim_{x \rightarrow 0} f(x)$. Since this limit is zero, $f(0) = 0$ is the only possible choice of $f(0)$ which makes f continuous at $x = 0$.

11. Substitution in Limits

Given two functions f and g one can consider their composition $h(x) = f(g(x))$. To compute the limit

$$\lim_{x \rightarrow a} f(g(x))$$

we write $u = g(x)$, so that we want to know

$$\lim_{x \rightarrow a} f(u) \text{ where } u = g(x).$$

Suppose that you can find the limits

$$L = \lim_{x \rightarrow a} g(x) \text{ and } \lim_{u \rightarrow L} f(u) = M.$$

Then it seems reasonable that as x approaches a , $u = g(x)$ will approach L , and $f(g(x))$ approaches M .

This is in fact a theorem:

11.1. Theorem. *If $\lim_{x \rightarrow a} g(x) = L$, and if the function f is continuous at $u = L$, then*

$$\lim_{x \rightarrow a} f(g(x)) = \lim_{u \rightarrow L} f(u) = f(L).$$

Another way to write this is

$$\lim_{x \rightarrow a} f(g(x)) = f\left(\lim_{x \rightarrow a} g(x)\right).$$

11.2. Example: compute $\lim_{x \rightarrow 3} \sqrt{x^3 - 3x^2 + 2}$. The given function is the composition of two functions, namely

$$\sqrt{x^3 - 3x^2 + 2} = \sqrt{u}, \text{ with } u = x^3 - 3x^2 + 2,$$

or, in function notation, we want to find $\lim_{x \rightarrow 3} h(x)$ where

$$h(x) = f(g(x)), \text{ with } g(x) = x^3 - 3x^2 + 2 \text{ and } f(u) = \sqrt{u}.$$

Either way, we have

$$\lim_{x \rightarrow 3} x^3 - 3x^2 + 2 = 2 \quad \text{and} \quad \lim_{u \rightarrow 2} \sqrt{u} = \sqrt{2}.$$

You get the first limit from the limit properties $(P_1) \dots (P_5)$. The second limit says that taking the square root is a continuous function, which it is. We have not proved that (yet), but this particular limit is the one from example 6.3. Putting these two limits together we conclude that the limit is $\sqrt{2}$.

Normally, you write this whole argument as follows:

$$\lim_{x \rightarrow 3} \sqrt{x^3 - 3x^2 + 2} = \sqrt{\lim_{x \rightarrow 3} x^3 - 3x^2 + 2} = \sqrt{2},$$

where you must point out that $f(x) = \sqrt{x}$ is a continuous function to justify the first step.

Another possible way of writing this is

$$\lim_{x \rightarrow 3} \sqrt{x^3 - 3x^2 + 2} = \lim_{u \rightarrow 2} \sqrt{u} = \sqrt{2},$$

where you must say that you have substituted $u = x^3 - 3x^2 + 2$.

12. Exercises

Find the following limits.

52. $\lim_{x \rightarrow -7} (2x + 5)$

53. $\lim_{x \rightarrow 7^-} (2x + 5)$

54. $\lim_{x \rightarrow -\infty} (2x + 5)$

55. $\lim_{x \rightarrow -4} (x + 3)^{2006}$

56. $\lim_{x \rightarrow -4} (x + 3)^{2007}$

57. $\lim_{x \rightarrow -\infty} (x + 3)^{2007}$

58. $\lim_{t \rightarrow 1} \frac{t^2 + t - 2}{t^2 - 1}$

59. $\lim_{t \nearrow 1} \frac{t^2 + t - 2}{t^2 - 1}$

60. $\lim_{t \rightarrow -1} \frac{t^2 + t - 2}{t^2 - 1}$

61. $\lim_{x \rightarrow \infty} \frac{x^2 + 3}{x^2 + 4}$

62. $\lim_{x \rightarrow \infty} \frac{x^5 + 3}{x^2 + 4}$

63. $\lim_{x \rightarrow \infty} \frac{x^2 + 1}{x^5 + 2}$

64. $\lim_{x \rightarrow \infty} \frac{(2x + 1)^4}{(3x^2 + 1)^2}$

65. $\lim_{u \rightarrow \infty} \frac{(2u + 1)^4}{(3u^2 + 1)^2}$

66. $\lim_{t \rightarrow 0} \frac{(2t + 1)^4}{(3t^2 + 1)^2}$

67. What are the coordinates of the points labeled A, \dots, E in Figure 2 (the graph of $y = \sin \pi/x$).

68. If $\lim_{x \rightarrow a} f(x)$ exists then f is continuous at $x = a$. *True or false?*

69. Give two examples of functions for which $\lim_{x \searrow 0} f(x)$ does not exist.

70. Group Problem.

If $\lim_{x \rightarrow 0} f(x)$ and $\lim_{x \rightarrow 0} g(x)$ both do not exist, then $\lim_{x \rightarrow 0} (f(x) + g(x))$ also does not exist. *True or false?*

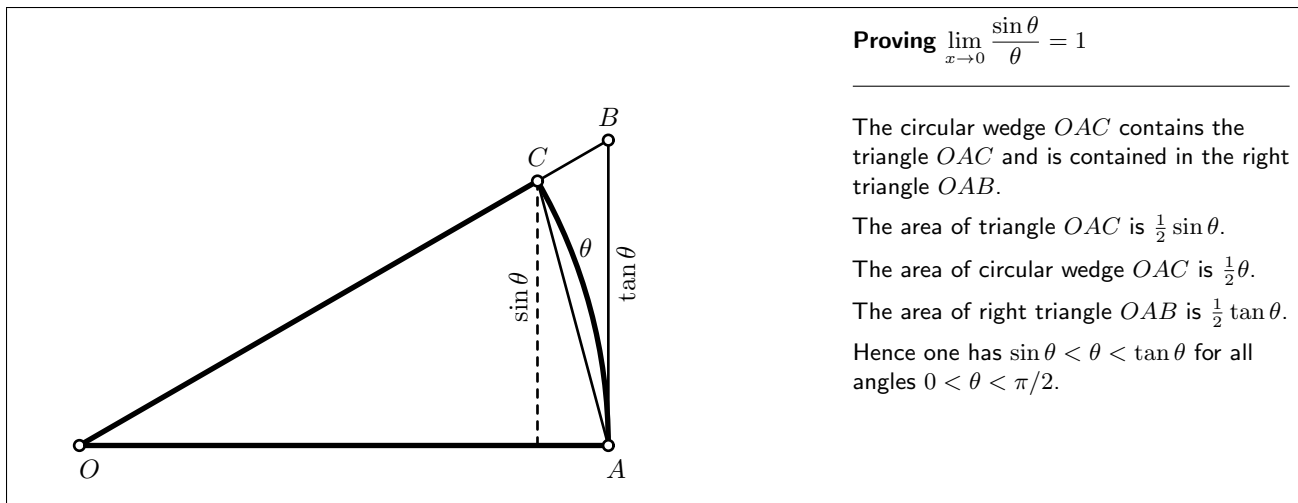
71. Group Problem.

If $\lim_{x \rightarrow 0} f(x)$ and $\lim_{x \rightarrow 0} g(x)$ both do not exist, then $\lim_{x \rightarrow 0} (f(x)/g(x))$ also does not exist. *True or false?*

72. Group Problem.

In the text we proved that $\lim_{x \rightarrow \infty} \frac{1}{x} = 0$. Show that this implies that $\lim_{x \rightarrow \infty} x$ does not exist. Hint: Suppose $\lim_{x \rightarrow \infty} x = L$ for some number L . Apply the limit properties to $\lim_{x \rightarrow \infty} x \cdot \frac{1}{x}$.

73. Evaluate $\lim_{x \rightarrow 9} \frac{\sqrt{x} - 3}{x - 9}$. Hint: Multiply top and bottom by $\sqrt{x} + 3$.



74. Evaluate $\lim_{x \rightarrow 2} \frac{\frac{1}{x} - \frac{1}{2}}{x - 2}$.

75. Evaluate $\lim_{x \rightarrow 2} \frac{\frac{1}{\sqrt{x}} - \frac{1}{\sqrt{2}}}{x - 2}$.

76. A function f is defined by

$$f(x) = \begin{cases} x^3 & \text{for } x < -1 \\ ax + b & \text{for } -1 \leq x < 1 \\ x^2 + 2 & \text{for } x \geq 1. \end{cases}$$

where a and b are constants. The function f is continuous. What are a and b ?

77. Find a constant k such that the function

$$f(x) = \begin{cases} 3x + 2 & \text{for } x < 2 \\ x^2 + k & \text{for } x \geq 2. \end{cases}$$

is continuous. Hint: Compute the one-sided limits.

78. Find constants a and c such that the function

$$f(x) = \begin{cases} x^3 + c & \text{for } x < 0 \\ ax + c^2 & \text{for } 0 \leq x < 1 \\ \arctan x & \text{for } x \geq 1. \end{cases}$$

is continuous for all x .

13. Two Limits in Trigonometry

In this section we'll derive a few limits involving the trigonometric functions. You can think of them as saying that for small angles θ one has

$$\sin \theta \approx \theta \quad \text{and} \quad \cos \theta \approx 1 - \frac{1}{2}\theta^2.$$

We will use these limits when we compute the derivatives of Sine, Cosine and Tangent.

13.1. Theorem. $\lim_{\theta \rightarrow 0} \frac{\sin \theta}{\theta} = 1.$

PROOF. The proof requires a few sandwiches and some geometry.

We begin by only considering positive angles, and in fact we will only consider angles $0 < \theta < \pi/2$.

Since the wedge OAC contains the triangle OAC its area must be larger. The area of the wedge is $\frac{1}{2}\theta$ and the area of the triangle is $\frac{1}{2}\sin \theta$, so we find that

$$(13) \quad 0 < \sin \theta < \theta \text{ for } 0 < \theta < \frac{\pi}{2}.$$

The Sandwich Theorem implies that

$$(14) \quad \lim_{\theta \searrow 0} \sin \theta = 0.$$

Moreover, we also have

$$(15) \quad \lim_{\theta \searrow 0} \cos \theta = \lim_{\theta \searrow 0} \sqrt{1 - \sin^2 \theta} = 1.$$

Next we compare the areas of the wedge OAC and the larger triangle OAB . Since OAB has area $\frac{1}{2} \tan \theta$ we find that

$$\theta < \tan \theta$$

for $0 < \theta < \frac{\pi}{2}$. Since $\tan \theta = \frac{\sin \theta}{\cos \theta}$ we can multiply with $\cos \theta$ and divide by θ to get

$$\cos \theta < \frac{\sin \theta}{\theta} \text{ for } 0 < \theta < \frac{\pi}{2}$$

If we go back to (15) and divide by θ , then we get

$$\cos \theta < \frac{\sin \theta}{\theta} < 1$$

The Sandwich Theorem can be used once again, and now it gives

$$\lim_{\theta \searrow 0} \frac{\sin \theta}{\theta} = 1.$$

This is a one-sided limit. To get the limit in which $\theta \nearrow 0$, you use that $\sin \theta$ is an odd function. □

13.2. An example. We will show that

$$(16) \quad \lim_{\theta \rightarrow 0} \frac{1 - \cos \theta}{\theta^2} = \frac{1}{2}.$$

This follows from $\sin^2 \theta + \cos^2 \theta = 1$. Namely,

$$\begin{aligned} \frac{1 - \cos \theta}{\theta^2} &= \frac{1}{1 + \cos \theta} \frac{1 - \cos^2 \theta}{\theta^2} \\ &= \frac{1}{1 + \cos \theta} \frac{\sin^2 \theta}{\theta^2} \\ &= \frac{1}{1 + \cos \theta} \left\{ \frac{\sin \theta}{\theta} \right\}^2. \end{aligned}$$

We have just shown that $\cos \theta \rightarrow 1$ and $\frac{\sin \theta}{\theta} \rightarrow 1$ as $\theta \rightarrow 0$, so (16) follows.

14. Exercises

Find each of the following limits or show that it does not exist. Distinguish between limits which are infinite and limits which do not exist.

79. $\lim_{\alpha \rightarrow 0} \frac{\sin 2\alpha}{\sin \alpha}$ (two ways: with and without the double angle formula!)

80. $\lim_{x \rightarrow 0} \frac{\sin 3x}{2x}$.

81. $\lim_{\theta \rightarrow 0} \frac{\tan \theta}{\theta}$.

82. $\lim_{\alpha \rightarrow 0} \frac{\tan 4\alpha}{\sin 2\alpha}$.

83. $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x \sin x}$.

84. $\lim_{\theta \rightarrow \pi/2} \frac{1 - \sin \theta}{\theta - \pi/2}$

85. $\lim_{x \rightarrow \infty} \frac{2x^3 + 3x^2 \cos x}{(x+2)^3}$.

86. $\lim_{x \rightarrow 0} \frac{\sin^2 x}{1 - \cos x}$.

87. $\lim_{x \rightarrow 0} \frac{\sin(x^2)}{x^2}$.

88. $\lim_{x \rightarrow 0} \frac{x(1 - \cos x)}{\tan^3 x}$.

89. $\lim_{x \rightarrow 0} \frac{\sin(x^2)}{1 - \cos x}$.

90. $\lim_{x \rightarrow \pi/2} \frac{x - \frac{\pi}{2}}{\cos x}$.

91. $\lim_{x \rightarrow \pi/2} (x - \frac{\pi}{2}) \tan x$.

92. $\lim_{x \rightarrow 0} \frac{\cos x}{x^2 + 9}$.

93. $\lim_{x \rightarrow \pi} \frac{\sin x}{x - \pi}$.

94. $\lim_{x \rightarrow 0} \frac{\sin x}{x + \sin x}$.

95. $A = \lim_{x \rightarrow \infty} \frac{\sin x}{x}$. **(!!)**

96. $B = \lim_{x \rightarrow \infty} \frac{\cos x}{x}$. **(!! again)**

97. Is there a constant k such that the function

$$f(x) = \begin{cases} \sin(1/x) & \text{for } x \neq 0 \\ k & \text{for } x = 0. \end{cases}$$

is continuous? If so, find it; if not, say why.

98. Find a constant A so that the function

$$f(x) = \begin{cases} \frac{\sin x}{2x} & \text{for } x \neq 0 \\ A & \text{when } x = 0 \end{cases}$$

99. Compute $\lim_{x \rightarrow \infty} x \sin \frac{\pi}{x}$ and $\lim_{x \rightarrow \infty} x \tan \frac{\pi}{x}$. (Hint: substitute something).

100. Group Problem.

(*Geometry & Trig review*) Let A_n be the area of the regular n -gon inscribed in the unit circle, and let B_n be the area of the regular n -gon whose inscribed circle has radius 1.

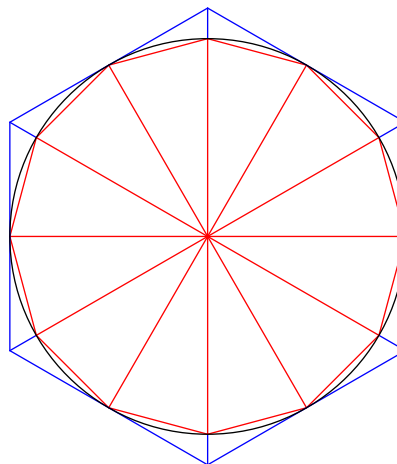
(a) Show that $A_n < \pi < B_n$.

(b) Show that

$$A_n = \frac{n}{2} \sin \frac{2\pi}{n} \text{ and } B_n = n \tan \frac{\pi}{n}$$

(c) Compute $\lim_{n \rightarrow \infty} A_n$ and $\lim_{n \rightarrow \infty} B_n$.

Here is a picture of A_{12} , B_6 and π :



On a historical note: Archimedes managed to compute A_{96} and B_{96} and by doing this got the most accurate approximation for π that was known in his time. See also:

http://www-history.mcs.st-andrews.ac.uk/HistTopics/Pi_through_the_ages.html

Derivatives (2)

“Leibniz never thought of the derivative as a limit”

<http://www.gap-system.org/~history/Biographies/Leibniz.html>

In chapter 2 we saw two mathematical problems which led to expressions of the form $\frac{0}{0}$. Now that we know how to handle limits, we can state the definition of the derivative of a function. After computing a few derivatives using the definition we will spend most of this section developing the *differential calculus*, which is a collection of rules that allow you to compute derivatives without always having to use basic definition.

1. Derivatives Defined

1.1. Definition. Let f be a function which is defined on some interval (c, d) and let a be some number in this interval.

The **derivative of the function f at a** is the value of the limit

$$(17) \quad f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}.$$

f is said to be **differentiable at a** if this limit exists.

f is called **differentiable on the interval (c, d)** if it is differentiable at every point a in (c, d) .

1.2. Other notations. One can substitute $x = a + h$ in the limit (17) and let $h \rightarrow 0$ instead of $x \rightarrow a$. This gives the formula

$$(18) \quad f'(a) = \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h},$$

Often you will find this equation written with x instead of a and Δx instead of h , which makes it look like this:

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

The interpretation is the same as in equation (6) from §4. The numerator $f(x + \Delta x) - f(x)$ represents the amount by which the function value of f changes if one increases its argument x by a (small) amount Δx . If you write $y = f(x)$ then we can call the increase in f

$$\Delta y = f(x + \Delta x) - f(x),$$

so that the derivative $f'(x)$ is

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}.$$

GOTTFRIED WILHELM VON LEIBNIZ, one of the inventors of calculus, came up with the idea that one should write this limit as

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x},$$

the idea being that after letting Δx go to zero it didn't vanish, but instead became an infinitely small quantity which Leibniz called “ dx .” The result of increasing x by this infinitely small quantity dx is that $y = f(x)$ increased by another infinitely small quantity dy . The ratio of these two infinitely small quantities is what we call the derivative of $y = f(x)$.

There are no “infinitely small real numbers,” and this makes Leibniz’ notation difficult to justify. In the 20th century mathematicians have managed to create a consistent theory of “infinitesimals” which allows you to compute with “ dx and dy ” as Leibniz and his contemporaries would have done. This theory is called “non standard analysis.” We won’t mention it any further¹. Nonetheless, even though we won’t use infinitely small numbers, Leibniz’ notation is very useful and we will use it.

2. Direct computation of derivatives

2.1. Example – The derivative of $f(x) = x^2$ is $f'(x) = 2x$. We have done this computation before in §2. The result was

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \rightarrow 0} (2x+h) = 2x.$$

Leibniz would have written

$$\frac{dx^2}{dx} = 2x.$$

2.2. The derivative of $g(x) = x$ is $g'(x) = 1$. Indeed, one has

$$g'(x) = \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} = \lim_{h \rightarrow 0} \frac{(x+h) - x}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1.$$

In Leibniz’ notation:

$$\frac{dx}{dx} = 1.$$

This is an example where Leibniz’ notation is most misleading, because if you divide dx by dx then you should of course get 1. Nonetheless, this is not what is going on. The expression $\frac{dx}{dx}$ is not really a fraction since there are no two “infinitely small” quantities dx which we are dividing.

2.3. The derivative of any constant function is zero . Let $k(x) = c$ be a constant function. Then we have

$$k'(x) = \lim_{h \rightarrow 0} \frac{k(x+h) - k(x)}{h} = \lim_{h \rightarrow 0} \frac{c - c}{h} = \lim_{h \rightarrow 0} 0 = 0.$$

Leibniz would have said that if c is a constant, then

$$\frac{dc}{dx} = 0.$$

2.4. Derivative of x^n for $n = 1, 2, 3, \dots$. To differentiate $f(x) = x^n$ one proceeds as follows:

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} \frac{x^n - a^n}{x - a}.$$

We need to simplify the fraction $(x^n - a^n)/(x - a)$. For $n = 2$ we have

$$\frac{x^2 - a^2}{x - a} = x + a.$$

For $n = 1, 2, 3, \dots$ the geometric sum formula tells us that

$$(19) \quad \frac{x^n - a^n}{x - a} = x^{n-1} + x^{n-2}a + x^{n-3}a^2 + \dots + xa^{n-2} + a^{n-1}.$$

¹But if you want to read more on this you should see Keisler’s calculus text at

<http://www.math.wisc.edu/~keisler/calc.html>

I would not recommend using Keisler’s text and this text at the same time, but if you like math you should remember that it exists, and look at it (later, say, after you pass 221.)

If you don't remember the geometric sum formula, then you could also just verify (19) by carefully multiplying both sides with $x - a$. For instance, when $n = 3$ you would get

$$\begin{array}{rcl} x \times (x^2 + xa + a^2) & = & x^3 + ax^2 + a^2x \\ -a \times (x^2 + xa + a^2) & = & -ax^2 - a^2x - a^3 \\ \hline (x - a) \times (x^2 + xa + a^2) & = & x^3 - a^3 \end{array}$$

With formula (19) in hand we can now easily find the derivative of x^n :

$$\begin{aligned} f'(a) &= \lim_{x \rightarrow a} \frac{x^n - a^n}{x - a} \\ &= \lim_{x \rightarrow a} \{x^{n-1} + x^{n-2}a + x^{n-3}a^2 + \cdots + xa^{n-2} + a^{n-1}\} \\ &= a^{n-1} + a^{n-2}a + a^{n-3}a^2 + \cdots + a^{n-2}a + a^{n-1}. \end{aligned}$$

Here there are n terms, and they all are equal to a^{n-1} , so the final result is

$$f'(a) = na^{n-1}.$$

One could also write this as $f'(x) = nx^{n-1}$, or, in Leibniz' notation

$$\frac{dx^n}{dx} = nx^{n-1}.$$

This formula turns out to be true in general, but here we have only proved it for the case in which n is a positive integer.

3. Differentiable implies Continuous

3.1. Theorem. *If a function f is differentiable at some a in its domain, then f is also continuous at a .*

PROOF. We are given that

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

exists, and we must show that

$$\lim_{x \rightarrow a} f(x) = f(a).$$

This follows from the following computation

$$\begin{aligned} \lim_{x \rightarrow a} f(x) &= \lim_{x \rightarrow a} (f(x) - f(a) + f(a)) && \text{(algebra)} \\ &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \cdot (x - a) + f(a) && \text{(more algebra)} \\ &= \left\{ \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \right\} \cdot \lim_{x \rightarrow a} (x - a) + \lim_{x \rightarrow a} f(a) && \text{(Limit Properties)} \\ &= f'(a) \cdot 0 + f(a) && (f'(a) \text{ exists}) \\ &= f(a). \end{aligned}$$

□

4. Some non-differentiable functions

4.1. A graph with a corner. Consider the function

$$f(x) = |x| = \begin{cases} x & \text{for } x \geq 0, \\ -x & \text{for } x < 0. \end{cases}$$

This function is continuous at all x , but it is not differentiable at $x = 0$.

To see this try to compute the derivative at 0,

$$f'(0) = \lim_{x \rightarrow 0} \frac{|x| - |0|}{x - 0} = \lim_{x \rightarrow 0} \frac{|x|}{x} = \lim_{x \rightarrow 0} \text{sign}(x).$$

We know this limit does not exist (see §7.2)

If you look at the graph of $f(x) = |x|$ then you see what is wrong: the graph has a corner at the origin and it is not clear which line, if any, deserves to be called the tangent to the graph at the origin.

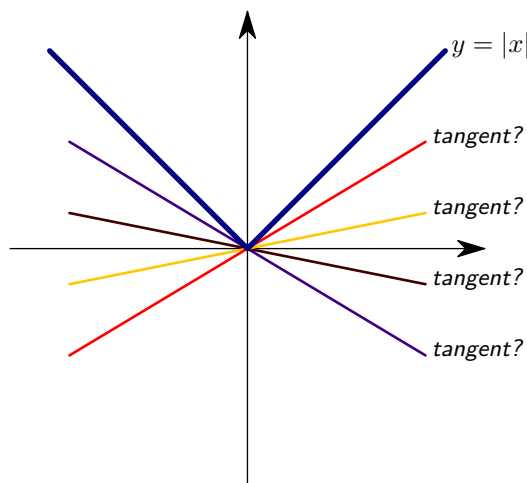


Figure 1. The graph of $y = |x|$ has no tangent at the origin.

4.2. A graph with a cusp. Another example of a function without a derivative at $x = 0$ is

$$f(x) = \sqrt{|x|}.$$

When you try to compute the derivative you get this limit

$$f'(0) = \lim_{x \rightarrow 0} \frac{\sqrt{|x|}}{x} = ?$$

The limit from the right is

$$\lim_{x \searrow 0} \frac{\sqrt{|x|}}{x} = \lim_{x \searrow 0} \frac{1}{\sqrt{x}},$$

which does not exist (it is $+\infty$). Likewise, the limit from the left also does not exist ('tis $-\infty$). Nonetheless, a drawing for the graph of f suggests an obvious tangent to the graph at $x = 0$, namely, the y -axis. That observation does not give us a derivative, because the y -axis is vertical and hence has no slope.

4.3. A graph with absolutely no tangents, *anywhere*. The previous two examples were about functions which did not have a derivative at $x = 0$. In both examples the point $x = 0$ was the only point where the function failed to have a derivative. It is easy to give examples of functions which are not differentiable at more than one value of x , but here I would like to show you a function f which doesn't have a derivative *anywhere in its domain*.

To keep things short I won't write a formula for the function, and merely show you a graph. In this graph you see a typical path of a Brownian motion, i.e. t is time, and $x(t)$ is the position of a particle which undergoes a Brownian motion – come to lecture for further explanation (see also the article on wikipedia). To see a similar graph check the Dow Jones or Nasdaq in the upper left hand corner of the web page at <http://finance.yahoo.com> in the afternoon on any weekday.

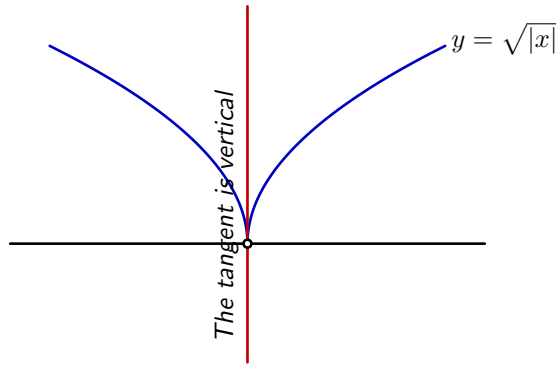


Figure 2. Tangent to the graph of $y = |x|^{1/2}$ at the origin

5. Exercises

101. Compute the derivative of the following functions

$$f(x) = x^2 - 2x$$

$$g(x) = \frac{1}{x}$$

$$k(x) = x^3 - 17x$$

$$u(x) = \frac{2}{1+x}$$

$$v(x) = \sqrt{x}$$

$$w(x) = \frac{1}{\sqrt{x}}$$

using either (17) or (18).

102. Which of the following functions is differentiable at $x = 0$?

$$f(x) = x|x|, \quad g(x) = x\sqrt{|x|},$$

$$h(x) = x + |x|, \quad k(x) = x^2 \sin \frac{\pi}{x},$$

$$\ell(x) = x \sin \frac{\pi}{x}.$$

These formulas do not define k and ℓ at $x = 0$. We define $k(0) = \ell(0) = 0$.

103. For which value(s) is the function defined by

$$f(x) = \begin{cases} ax + b & \text{for } x < 0 \\ x - x^2 & \text{for } x \geq 0 \end{cases}$$

differentiable at $x = 0$? Sketch the graph of the function f for the values a and b you found.

104. For which value(s) is the function defined by

$$f(x) = \begin{cases} ax^2 + b & \text{for } x < 1 \\ x - x^2 & \text{for } x \geq 1 \end{cases}$$

differentiable at $x = 0$? Sketch the graph of the function f for the values a and b you found.

105. For which value(s) is the function defined by

$$f(x) = \begin{cases} ax^2 & \text{for } x < 2 \\ x + b & \text{for } x \geq 2 \end{cases}$$

differentiable at $x = 0$? Sketch the graph of the function f for the values a and b you found.

106. Group Problem.

True or false: If a function f is continuous at some $x = a$ then it must also be differentiable at $x = a$?

107. Group Problem.

True or false: If a function f is differentiable at some $x = a$ then it must also be continuous at $x = a$?

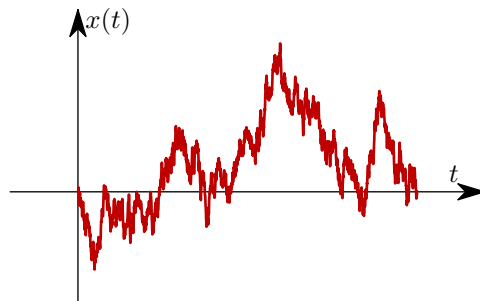


Figure 3. A Brownian motion. Note how the graph doesn't seem to have a tangent anywhere at all.

6. The Differentiation Rules

You could go on and compute more derivatives from the definition. Each time you would have to compute a new limit, and hope that there is some trick that allows you to find that limit. This is fortunately not necessary. It turns out that if you know a few basic derivatives (such as $dx^n/dx = nx^{n-1}$) the you can find derivatives of arbitrarily complicated functions by breaking them into smaller pieces. In this section we'll look at rules which tell you how to differentiate a function which is either the sum, difference, product or quotient of two other functions.

| | | |
|-----------------------|---|---|
| <i>Constant rule:</i> | $c' = 0$ | $\frac{dc}{dx} = 0$ |
| <i>Sum rule:</i> | $(u \pm v)' = u' \pm v'$ | $\frac{du \pm v}{dx} = \frac{du}{dx} \pm \frac{dv}{dx}$ |
| <i>Product rule:</i> | $(u \cdot v)' = u' \cdot v + u \cdot v'$ | $\frac{d(uv)}{dx} = \frac{du}{dx}v + u\frac{dv}{dx}$ |
| <i>Quotient rule:</i> | $\left(\frac{u}{v}\right)' = \frac{u' \cdot v - u \cdot v'}{v^2}$ | $\frac{d\frac{u}{v}}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$ |

Table 1. The differentiation rules

The situation is analogous to that of the “limit-properties” $(P_1) \dots (P_6)$ from the previous chapter which allowed us to compute limits without always having to go back to the epsilon-delta definition.

6.1. Sum, product and quotient rules. In the following c and n are constants, u and v are functions of x , and $'$ denotes differentiation. The Differentiation Rules in function notation, and Leibniz notation, are listed in figure 1.

Note that we already proved the Constant Rule in example 2.2. We will now prove the sum, product and quotient rules.

6.2. Proof of the Sum Rule. Suppose that $f(x) = u(x) + v(x)$ for all x where u and v are differentiable. Then

$$\begin{aligned}
 f'(a) &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} && \text{(definition of } f') \\
 &= \lim_{x \rightarrow a} \frac{(u(x) + v(x)) - (u(a) + v(a))}{x - a} && \text{(use } f = u + v) \\
 &= \lim_{x \rightarrow a} \left(\frac{u(x) - u(a)}{x - a} + \frac{v(x) - v(a)}{x - a} \right) && \text{(algebra)} \\
 &= \lim_{x \rightarrow a} \frac{u(x) - u(a)}{x - a} + \lim_{x \rightarrow a} \frac{v(x) - v(a)}{x - a} && \text{(limit property)} \\
 &= u'(a) + v'(a) && \text{(definition of } u', v')
 \end{aligned}$$

6.3. Proof of the Product Rule. Let $f(x) = u(x)v(x)$. To find the derivative we must express the change of f in terms of the changes of u and v

$$\begin{aligned}
 f(x) - f(a) &= u(x)v(x) - u(a)v(a) \\
 &= u(x)v(x) - u(x)v(a) + u(x)v(a) - u(a)v(a) \\
 &= u(x)(v(x) - v(a)) + (u(x) - u(a))v(a)
 \end{aligned}$$

Now divide by $x - a$ and let $x \rightarrow a$:

$$\begin{aligned}\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} &= \lim_{x \rightarrow a} u(x) \frac{v(x) - v(a)}{x - a} + \frac{u(x) - u(a)}{x - a} v(a) \\ &\quad \text{(use the limit properties)} \\ &= \left(\lim_{x \rightarrow a} u(x) \right) \left(\lim_{x \rightarrow a} \frac{v(x) - v(a)}{x - a} \right) + \left(\lim_{x \rightarrow a} \frac{u(x) - u(a)}{x - a} \right) v(a) \\ &= u(a)v'(a) + u'(a)v(a),\end{aligned}$$

as claimed. In this last step we have used that

$$\lim_{x \rightarrow a} \frac{u(x) - u(a)}{x - a} = u'(a) \quad \text{and} \quad \lim_{x \rightarrow a} \frac{v(x) - v(a)}{x - a} = v'(a)$$

and also that

$$\lim_{x \rightarrow a} u(x) = u(a)$$

This last limit follows from the fact that u is continuous, which in turn follows from the fact that u is differentiable.

6.4. Proof of the Quotient Rule . We can break the proof into two parts. First we do the special case where $f(x) = 1/v(x)$, and then we use the product rule to differentiate

$$f(x) = \frac{u(x)}{v(x)} = u(x) \cdot \frac{1}{v(x)}.$$

So let $f(x) = 1/v(x)$. We can express the change in f in terms of the change in v

$$f(x) - f(a) = \frac{1}{v(x)} - \frac{1}{v(a)} = \frac{v(x) - v(a)}{v(x)v(a)}.$$

Dividing by $x - a$ we get

$$\frac{f(x) - f(a)}{x - a} = \frac{1}{v(x)v(a)} \frac{v(x) - v(a)}{x - a}.$$

Now we want to take the limit $x \rightarrow a$. We are given the v is differentiable, so it must also be continuous and hence

$$\lim_{x \rightarrow a} v(x) = v(a).$$

Therefore we find

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} \frac{1}{v(x)v(a)} \lim_{x \rightarrow a} \frac{v(x) - v(a)}{x - a} = \frac{v'(a)}{v(a)^2}.$$

That completes the first step of the proof. In the second step we use the product rule to differentiate $f = u/v$

$$f' = \left(\frac{u}{v} \right)' = \left(u \cdot \frac{1}{v} \right)' = u' \cdot \frac{1}{v} + u \cdot \left(\frac{1}{v} \right)' = \frac{u'}{v} - u \frac{v'}{v^2} = \frac{u'v - uv'}{v^2}.$$

6.5. A shorter, but not quite perfect derivation of the Quotient Rule . The Quotient Rule can be derived from the Product Rule as follows: if $w = u/v$ then

$$(20) \quad w \cdot v = u$$

By the product rule we have

$$w' \cdot v + w \cdot v' = u',$$

so that

$$w' = \frac{u' - w \cdot v'}{v} = \frac{u' - (u/v) \cdot v'}{v} = \frac{u' \cdot v - u \cdot v'}{v^2}.$$

Unlike the proof in §6.4 above, this argument does not prove that w is differentiable if u and v are. It only says that *if the derivative exists* then it must be what the Quotient Rule says it is.

The trick which is used here, is a special case of a method called “implicit differentiation.” We have an equation (20) which the quotient w satisfies, and from by differentiating this equation we find w' .

6.6. Differentiating a constant multiple of a function . Note that the rule

$$(cu)' = cu'$$

follows from the Constant Rule and the Product Rule.

6.7. Picture of the Product Rule. If u and v are quantities which depend on x , and if increasing x by Δx causes u and v to change by Δu and Δv , then the product of u and v will change by

$$(21) \quad \Delta(uv) = (u + \Delta u)(v + \Delta v) - uv = u\Delta v + v\Delta u + \Delta u\Delta v.$$

If u and v are differentiable functions of x , then the changes Δu and Δv will be of the same order of magnitude as Δx , and thus one expects $\Delta u\Delta v$ to be much smaller. One therefore ignores the last term in (21), and thus arrives at

$$\Delta(uv) = u\Delta v + v\Delta u.$$

Leibniz would now divide by Δx and replace Δ 's by d 's to get the product rule:

$$\frac{\Delta(uv)}{\Delta x} = u \frac{\Delta v}{\Delta x} + v \frac{\Delta u}{\Delta x}.$$

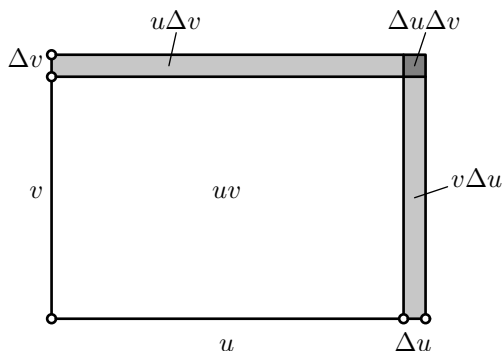


Figure 4. The Product Rule. *How much does the area of a rectangle change if its sides u and v are increased by Δu and Δv ? Most of the increase is accounted for by the two thin rectangles whose areas are $u\Delta v$ and $v\Delta u$. So the increase in area is approximately $u\Delta v + v\Delta u$, which explains why the product rule says $(uv)' = uv' + vu'$.*

7. Differentiating powers of functions

7.1. Product rule with more than one factor. If a function is given as the product of n functions, i.e.

$$f(x) = u_1(x) \times u_2(x) \times \cdots \times u_n(x),$$

then you can differentiate it by applying the product rule $n - 1$ times (there are n factors, so there are $n - 1$ multiplications.)

After the first step you would get

$$f' = u_1'(u_2 \cdots u_n) + u_1(u_2 \cdots u_n)'$$

In the second step you apply the product rule to $(u_2 u_3 \cdots u_n)'$. This yields

$$\begin{aligned} f' &= u_1' u_2 \cdots u_n + u_1 [u_2' u_3 \cdots u_n + u_2 (u_3 \cdots u_n)'] \\ &= u_1' u_2 \cdots u_n + u_1 u_2' u_3 \cdots u_n + u_1 u_2 (u_3 \cdots u_n)'. \end{aligned}$$

Continuing this way one finds after $n - 1$ applications of the product rule that

$$(22) \quad (u_1 \cdots u_n)' = u_1' u_2 \cdots u_n + u_1 u_2' u_3 \cdots u_n + \cdots + u_1 u_2 u_3 \cdots u_n'.$$

7.2. The Power rule . If all n factors in the previous paragraph are the same, so that the function f is the n^{th} power of some other function,

$$f(x) = (u(x))^n,$$

then all terms in the right hand side of (22) are the same, and, since there are n of them, one gets

$$f'(x) = nu^{n-1}(x)u'(x),$$

or, in Leibniz' notation,

$$(23) \quad \frac{du^n}{dx} = nu^{n-1} \frac{du}{dx}.$$

7.3. The Power Rule for Negative Integer Exponents . We have just proved the power rule (23) assuming n is a positive integer. The rule actually holds for all real exponents n , but the proof is harder.

Here we prove the Power Rule for negative exponents using the Quotient Rule. Suppose $n = -m$ where m is a positive integer. Then the Quotient Rule tells us that

$$(u^n)' = (u^{-m})' = \left(\frac{1}{u^m} \right)' \stackrel{\text{Q.R.}}{=} -\frac{(u^m)'}{(u^m)^2}.$$

Since m is a positive integer, we can use (23), so $(u^m)' = mu^{m-1}$, and hence

$$(u^n)' = -\frac{mu^{m-1} \cdot u'}{u^{2m}} = -mu^{-m-1} \cdot u' = nu^{n-1}u'.$$

7.4. The Power Rule for Rational Exponents . So far we have proved that the power law holds if the exponent n is an integer.

We will now see how you can show that the power law holds even if the exponent n is any fraction, $n = p/q$. The following derivation contains the trick called *implicit differentiation* which we will study in more detail in Section 15.

So let $n = p/q$ where p and q are integers and consider the function

$$w(x) = u(x)^{p/q}.$$

Assuming that both u and w are differentiable functions, we will show that

$$(24) \quad w'(x) = \frac{p}{q} u(x)^{\frac{p}{q}-1} u'(x)$$

Raising both sides to the q th power gives

$$w(x)^q = u(x)^p.$$

Here the exponents p and q are integers, so we may apply the Power Rule to both sides. We get

$$qw^{q-1} \cdot w' = pu^{p-1} \cdot u'.$$

Dividing both sides by qw^{q-1} and substituting $u^{p/q}$ for w gives

$$w' = \frac{pu^{p-1} \cdot u'}{qw^{q-1}} = \frac{pu^{p-1} \cdot u'}{qu^{p(q-1)/q}} = \frac{pu^{p-1} \cdot u'}{qu^{p-(p/q)}} = \frac{p}{q} \cdot u^{(p/q)-1} \cdot u'$$

which is the Power Rule for $n = p/q$.

This proof is flawed because we did not show that $w(x) = u(x)^{p/q}$ is differentiable: we only showed what the derivative should be, *if it exists*.

7.5. Derivative of x^n for integer n . If you choose the function $u(x)$ in the Power Rule to be $u(x) = x$, then $u'(x) = 1$, and hence the derivative of $f(x) = u(x)^n = x^n$ is

$$f'(x) = nu(x)^{n-1}u'(x) = nx^{n-1} \cdot 1 = nx^{n-1}.$$

We already knew this of course.

7.6. Example – differentiate a polynomial. Using the Differentiation Rules you can easily differentiate any polynomial and hence any rational function. For example, using the Sum Rule, the Power Rule with $u(x) = x$, the rule $(cu)' = cu'$, the derivative of the polynomial

$$f(x) = 2x^4 - x^3 + 7$$

is

$$f'(x) = 8x^3 - 3x^2.$$

7.7. Example – differentiate a rational function. By the Quotient Rule the derivative of the function

$$g(x) = \frac{2x^4 - x^3 + 7}{1 + x^2}$$

is

$$\begin{aligned} g'(x) &= \frac{(8x^3 - 3x^2)(1 + x^2) - (2x^4 - x^3 + 7)2x}{(1 + x^2)^2} \\ &= \frac{6x^5 - x^4 + 8x^3 - 3x^2 - 14x}{(1 + x^2)^2}. \end{aligned}$$

If you compare this example with the previous then you see that polynomials simplify when you differentiate them while rational functions become more complicated.

7.8. Derivative of the square root. The derivative of $f(x) = \sqrt{x} = x^{1/2}$ is

$$f'(x) = \frac{1}{2}x^{1/2-1} = \frac{1}{2}x^{-1/2} = \frac{1}{2x^{1/2}} = \frac{1}{2\sqrt{x}}$$

where we used the power rule with $n = 1/2$ and $u(x) = x$.

8. Exercises

108. Let $f(x) = (x^2 + 1)(x^3 + 3)$. Find $f'(x)$ in two ways:

(a) by multiplying and then differentiating,

(b) by using the product rule.

Are your answers the same?

109. Let $f(x) = (1 + x^2)^4$. Find $f'(x)$ in two ways, first by expanding to get an expression for $f(x)$ as a polynomial in x and then differentiating, and then by using the power rule. Are the answers the same?

110. Prove the statement in §6.6, i.e. show that $(cu)' = c(u')$ follows from the product rule.

Compute the derivatives of the following functions.

(try to simplify your answers)

111. $f(x) = x + 1 + (x + 1)^2$

112. $f(x) = \frac{x - 2}{x^4 + 1}$

113. $f(x) = \left(\frac{1}{1 + x}\right)^{-1}$

114. $f(x) = \sqrt{1 - x^2}$

115. $f(x) = \frac{ax + b}{cx + d}$

116. $f(x) = \frac{1}{(1 + x^2)^2}$

117. $f(x) = \frac{x}{1 + \sqrt{x}}$

118. $f(x) = \sqrt{\frac{1 - x}{1 + x}}$

119. $f(x) = \sqrt[3]{x + \sqrt{x}}$

120. $\varphi(t) = \frac{t}{1 + \sqrt{t}}$

121. $g(s) = \sqrt{\frac{1 - s}{1 + s}}$

122. $h(\rho) = \sqrt[3]{\rho + \sqrt{\rho}}$

123. Group Problem.

Using derivatives to approximate numbers.

(a) Find the derivative of $f(x) = x^{4/3}$.

(b) Use (a) to estimate the number

$$\frac{127^{4/3} - 125^{4/3}}{2}$$

approximately without a calculator. Your answer should have the form p/q where p and q are integers. [Hint: Note that $5^3 = 125$ and take a good look at equation (17).]

(c) Approximate in the same way the numbers $\sqrt{143}$ and $\sqrt{145}$ (Hint: $12 \times 12 = 144$).

124. Group Problem.

(Making the product and quotient rules look nicer.)

Instead of looking at the derivative of a function you can look at the ratio of its derivative to the function itself, i.e. you can compute f'/f . This quantity is called the **logarithmic derivative of the function** f for reasons that will become clear later this semester.

(a) Compute the logarithmic derivative of these functions (i.e. find $f'(x)/f(x)$)

$$F(x) = x, \quad g(x) = 3x, \quad h(x) = x^2 \\ k(x) = -x^2, \quad \ell(x) = 2007x^2, \quad m(x) = x^{2007}$$

(b) Show that for any pair of functions u and v one has

$$\frac{(uv)'}{uv} = \frac{u'}{u} + \frac{v'}{v} \\ \frac{(u/v)'}{u/v} = \frac{u'}{u} - \frac{v'}{v} \\ \frac{(u^n)'}{u^n} = n \frac{u'}{u}$$

125. (a) Find $f'(x)$ and $g'(x)$ if

$$f(x) = \frac{1+x^2}{2x^4+7}, \quad g(x) = \frac{2x^4+7}{1+x^2}.$$

Note that $f(x) = 1/g(x)$.

(b) Is it true that $f'(x) = 1/g'(x)$?

(c) Is it true that $f(x) = g^{-1}(x)$?

(d) Is it true that $f(x) = g(x)^{-1}$?

126. Group Problem. (a) Let $x(t) = (1-t^2)/(1+t^2)$, $y(t) = 2t/(1+t^2)$ and $u(t) = y(t)/x(t)$. Find dx/dt , dy/dt .

(b) Now that you've done (a) there are two different ways of finding du/dt . What are they, and use one of both to find du/dt .

9. Higher Derivatives

9.1. The derivative is a function. If the derivative $f'(a)$ of some function f exists for all a in the domain of f , then we have a new function: namely, for each number in the domain of f we compute the derivative of f at that number. This function is called the **derivative function** of f , and it is denoted by f' . Now that we have agreed that the derivative of a function is a function, we can repeat the process and try to differentiate the derivative. The result, if it exists, is called the **second derivative of f** . It is denoted f'' . The derivative of the second derivative is called the third derivative, written f''' , and so on.

The n th derivative of f is denoted $f^{(n)}$. Thus

$$f^{(0)} = f, \quad f^{(1)} = f', \quad f^{(2)} = f'', \quad f^{(3)} = f''', \dots$$

Leibniz' notation for the n th derivative of $y = f(x)$ is

$$\frac{d^n y}{dx^n} = f^{(n)}(x).$$

9.2. Example. If $f(x) = x^2 - 2x + 3$ then

$$f(x) = x^2 - 2x + 3 \\ f'(x) = 2x - 2 \\ f''(x) = 2 \\ f^{(3)}(x) = 0 \\ f^{(4)}(x) = 0 \\ \vdots$$

All further derivatives of f are zero.

9.3. Operator notation. A common variation on Leibniz' notation for derivatives is the so-called *operator notation*, as in

$$\frac{d(x^3 - x)}{dx} = \frac{d}{dx}(x^3 - x) = 3x^2 - 1.$$

For higher derivatives one can write

$$\frac{d^2 y}{dx^2} = \left(\frac{d}{dx} \right)^2 y$$

Be careful to distinguish the second derivative from the square of the first derivative. Usually

$$\frac{d^2 y}{dx^2} \neq \left(\frac{dy}{dx} \right)^2 \text{ !!!!}$$

10. Exercises

127. The equation

$$(\dagger) \quad \frac{2x}{x^2 - 1} = \frac{1}{x + 1} + \frac{1}{x - 1}$$

holds for all values of x (except $x = \pm 1$), so you should get the same answer if you differentiate both sides. Check this.

Compute the third derivative of $f(x) = 2x/(x^2 - 1)$ by using either the left or right hand side (your choice) of (\dagger) .

128. Compute the first, second and third derivatives of the following functions

$$f(x) = (x + 1)^4$$

$$g(x) = (x^2 + 1)^4$$

$$h(x) = \sqrt{x - 2}$$

$$k(x) = \sqrt[3]{x - \frac{1}{x}}$$

129. Find the derivatives of 10th order of the functions

$$f(x) = x^{12} + x^8$$

$$g(x) = \frac{1}{x}$$

$$h(x) = \frac{12}{1 - x}$$

$$k(x) = \frac{x^2}{1 - x}$$

130. Find $f'(x)$, $f''(x)$ and $f^{(3)}(x)$ if

$$f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720}.$$

131. Group Problem.

(a) Find the 12th derivative of the function $f(x) = \frac{1}{x + 2}$.

(b) Find the n^{th} order derivative of $f(x) = \frac{1}{x + 2}$ (i.e. find a formula for $f^{(n)}(x)$ which is valid for all $n = 0, 1, 2, 3, \dots$).

(c) Find the n^{th} order derivative of $g(x) = \frac{x}{x + 2}$.

132. (About notation.)

(a) Find dy/dx and d^2y/dx^2 if $y = x/(x + 2)$.

(b) Find du/dt and d^2u/dt^2 if $u = t/(t + 2)$. Hint: See previous problem.

(c) Find $\frac{d}{dx} \left(\frac{x}{x + 2} \right)$ and $\frac{d^2}{dx^2} \left(\frac{x}{x + 2} \right)$. Hint: See previous problem.

(d) Find $\frac{d}{dx} \left(\frac{x}{x + 2} \right) \Big|_{x=1}$ and $\frac{d}{dx} \left(\frac{1}{1 + 2} \right)$.

133. Find d^2y/dx^2 and $(dy/dx)^2$ if $y = x^3$.

11. Differentiating Trigonometric functions

The trigonometric functions Sine, Cosine and Tangent are differentiable, and their derivatives are given by the following formulas

$$(25) \quad \frac{d \sin x}{dx} = \cos x, \quad \frac{d \cos x}{dx} = -\sin x, \quad \frac{d \tan x}{dx} = \frac{1}{\cos^2 x}.$$

Note the minus sign in the derivative of the cosine!

PROOF. By definition one has

$$\sin'(x) = \lim_{h \rightarrow 0} \frac{\sin(x + h) - \sin(x)}{h}.$$

To simplify the numerator we use the trigonometric addition formula

$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

with $\alpha = x$ and $\beta = h$, which results in

$$\begin{aligned}\frac{\sin(x+h) - \sin(x)}{h} &= \frac{\sin(x)\cos(h) + \cos(x)\sin(h) - \sin(x)}{h} \\ &= \cos(x)\frac{\sin(h)}{h} + \sin(x)\frac{\cos(h) - 1}{h}\end{aligned}$$

Hence by the formulas

$$\lim_{h \rightarrow 0} \frac{\sin(h)}{h} = 1 \quad \text{and} \quad \lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h} = 0$$

from Section 13 we have

$$\begin{aligned}\sin'(x) &= \lim_{h \rightarrow 0} \cos(x)\frac{\sin(h)}{h} + \sin(x)\frac{\cos(h) - 1}{h} \\ &= \cos(x) \cdot 1 + \sin(x) \cdot 0 \\ &= \cos(x).\end{aligned}$$

A similar computation leads to the stated derivative of $\cos x$.

To find the derivative of $\tan x$ we apply the quotient rule to

$$\tan x = \frac{\sin x}{\cos x} = \frac{f(x)}{g(x)}.$$

We get

$$\tan'(x) = \frac{\cos(x)\sin'(x) - \sin(x)\cos'(x)}{\cos^2(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}$$

as claimed. □

12. Exercises

Find the derivatives of the following functions (try to simplify your answers)

134. $f(x) = \sin(x) + \cos(x)$

135. $f(x) = 2\sin(x) - 3\cos(x)$

136. $f(x) = 3\sin(x) + 2\cos(x)$

137. $f(x) = x\sin(x) + \cos(x)$

138. $f(x) = x\cos(x) - \sin x$

139. $f(x) = \frac{\sin x}{x}$

140. $f(x) = \cos^2(x)$

141. $f(x) = \sqrt{1 - \sin^2 x}$

142. $f(x) = \sqrt{\frac{1 - \sin x}{1 + \sin x}}$

143. $\cot(x) = \frac{\cos x}{\sin x}$.

144. Can you find a and b so that the function

$$f(x) = \begin{cases} \cos x & \text{for } x \leq \frac{\pi}{4} \\ a + bx & \text{for } x > \frac{\pi}{4} \end{cases}$$

is differentiable at $x = \pi/4$?

145. Can you find a and b so that the function

$$f(x) = \begin{cases} \tan x & \text{for } x < \frac{\pi}{6} \\ a + bx & \text{for } x \geq \frac{\pi}{6} \end{cases}$$

is differentiable at $x = \pi/4$?

146. If f is a given function, and you have another function g which satisfies $g(x) = f(x) + 12$ for all x , then f and g have the same derivatives. Prove this. [*Hint: it's a short proof – use the differentiation rules.*]

147. **Group Problem.**

Show that the functions

$$f(x) = \sin^2 x \text{ and } g(x) = -\cos^2 x$$

have the same derivative by computing $f'(x)$ and $g'(x)$.

With hindsight this was to be expected – why?

148. Find the first and second derivatives of the functions

$$f(x) = \tan^2 x \text{ and } g(x) = \frac{1}{\cos^2 x}.$$

Hint: remember your trig to reduce work!

A depends on B depends on C depends on . . .

Someone is pumping water into a balloon. Assuming that the balloon is spherical you can say how large it is by specifying its radius R . For a growing balloon this radius will change with time t .

The volume of the balloon is a function of its radius, since the volume of a sphere of radius r is given by

$$V = \frac{4}{3}\pi r^3.$$

We now have two functions, the first f turns tells you the radius r of the balloon at time t ,

$$r = f(t)$$

and the second tells you the volume of the balloon given its radius

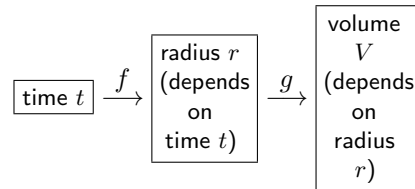
$$V = g(r).$$

The volume of the balloon at time t is then given by

$$V = g(f(t)) = g \circ f(t),$$

i.e. the function which tells you the volume of the balloon at time t is the composition of first f and then g .

Schematically we can summarize this chain of cause-and-effect relations as follows: you could either say that V depends on r , and r depends on t ,



or you could say that V depends directly on t :

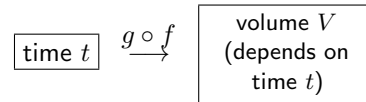


Figure 5. A “real world example” of a composition of functions.

13. The Chain Rule

13.1. Composition of functions. Given two functions f and g , one can define a new function called the **composition of f and g** . The notation for the composition is $f \circ g$, and it is defined by the formula

$$f \circ g(x) = f(g(x)).$$

The domain of the composition is the set of all numbers x for which this formula gives you something well-defined.

For instance, if $f(x) = x^2 + x$ and $g(x) = 2x + 1$ then

$$f \circ g(x) = f(2x + 1) = (2x + 1)^2 + (2x + 1)$$

$$\text{and } g \circ f(x) = g(x^2 + x) = 2(x^2 + x) + 1$$

Note that $f \circ g$ and $g \circ f$ are not the same function in this example (they hardly ever are the same).

If you think of functions as expressing dependence of one quantity on another, then the composition of functions arises as follows. If a quantity z is a function of another quantity y , and if y itself depends on x , then z depends on x via y .

To get $f \circ g$ from the previous example, we could say $z = f(y)$ and $y = g(x)$, so that

$$z = f(y) = y^2 + y \text{ and } y = 2x + 1.$$

Given x one can compute y , and from y one can then compute z . The result will be

$$z = y^2 + y = (2x + 1)^2 + (2x + 1),$$

in other notation,

$$z = f(y) = f(g(x)) = f \circ g(x).$$

One says that **the composition of f and g is the result of substituting g in f** .

13.2. Theorem (Chain Rule). *If f and g are differentiable, so is the composition $f \circ g$.*

The derivative of $f \circ g$ is given by

$$(f \circ g)'(x) = f'(g(x)) g'(x).$$

The chain rule tells you how to find the derivative of the composition $f \circ g$ of two functions f and g provided you now how to differentiate the two functions f and g .

When written in Leibniz' notation the chain rule looks particularly easy. Suppose that $y = g(x)$ and $z = f(y)$, then $z = f \circ g(x)$, and the derivative of z with respect to x is the derivative of the function $f \circ g$. The derivative of z with respect to y is the derivative of the function f , and the derivative of y with respect to x is the derivative of the function g . In short,

$$\frac{dz}{dx} = (f \circ g)'(x), \quad \frac{dz}{dy} = f'(y) \quad \text{and} \quad \frac{dy}{dx} = g'(x)$$

so that the chain rule says

$$(26) \quad \frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}.$$

FIRST PROOF OF THE CHAIN RULE (USING LEIBNIZ' NOTATION). We first consider difference quotients instead of derivatives, i.e. using the same notation as above, we consider the effect of an increase of x by an amount Δx on the quantity z .

If x increases by Δx , then $y = g(x)$ will increase by

$$\Delta y = g(x + \Delta x) - g(x),$$

and $z = f(y)$ will increase by

$$\Delta z = f(y + \Delta y) - f(y).$$

The ratio of the increase in $z = f(g(x))$ to the increase in x is

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \cdot \frac{\Delta y}{\Delta x}.$$

In contrast to dx , dy and dz in equation (26), the Δx , etc. here are finite quantities, so this equation is just algebra: you can cancel the two Δy s. If you let the increase Δx go to zero, then the increase Δy will also go to zero, and the difference quotients converge to the derivatives,

$$\frac{\Delta z}{\Delta x} \longrightarrow \frac{dz}{dx}, \quad \frac{\Delta z}{\Delta y} \longrightarrow \frac{dz}{dy}, \quad \frac{\Delta y}{\Delta x} \longrightarrow \frac{dy}{dx}$$

which immediately leads to Leibniz' form of the quotient rule. □

PROOF OF THE CHAIN RULE. We verify the formula in Theorem 13.2 at some arbitrary value $x = a$, i.e. we will show that

$$(f \circ g)'(a) = f'(g(a)) g'(a).$$

By definition the left hand side is

$$(f \circ g)'(a) = \lim_{x \rightarrow a} \frac{(f \circ g)(x) - (f \circ g)(a)}{x - a} = \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{x - a}.$$

The two derivatives on the right hand side are given by

$$g'(a) = \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a}$$

and

$$f'(g(a)) = \lim_{y \rightarrow a} \frac{f(y) - f(g(a))}{y - g(a)}.$$

Since g is a differentiable function it must also be a continuous function, and hence $\lim_{x \rightarrow a} g(x) = g(a)$. So we can substitute $y = g(x)$ in the limit defining $f'(g(a))$

$$(27) \quad f'(g(a)) = \lim_{y \rightarrow a} \frac{f(y) - f(g(a))}{y - g(a)} = \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)}.$$

Put all this together and you get

$$\begin{aligned} (f \circ g)'(a) &= \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{x - a} \\ &= \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \frac{g(x) - g(a)}{x - a} \\ &= \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \cdot \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \\ &= f'(g(a)) \cdot g'(a) \end{aligned}$$

which is what we were supposed to prove – the proof seems complete.

There is one flaw in this proof, namely, we have divided by $g(x) - g(a)$, which is not allowed when $g(x) - g(a) = 0$. This flaw can be fixed but we will not go into the details here.² \square

13.3. First example. We go back to the functions

$$z = f(y) = y^2 + y \text{ and } y = g(x) = 2x + 1$$

from the beginning of this section. The composition of these two functions is

$$z = f(g(x)) = (2x + 1)^2 + (2x + 1) = 4x^2 + 6x + 2.$$

We can compute the derivative of this composed function, i.e. the derivative of z with respect to x in two ways. First, you simply differentiate the last formula we have:

$$(28) \quad \frac{dz}{dx} = \frac{d(4x^2 + 6x + 2)}{dx} = 8x + 6.$$

The other approach is to use the chain rule:

$$\frac{dz}{dy} = \frac{d(y^2 + y)}{dy} = 2y + 1,$$

and

$$\frac{dy}{dx} = \frac{d(2x + 1)}{dx} = 2.$$

Hence, by the chain rule one has

$$(29) \quad \frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx} = (2y + 1) \cdot 2 = 4y + 2.$$

The two answers (28) and (29) should be the same. Once you remember that $y = 2x + 1$ you see that this is indeed true:

$$y = 2x + 1 \implies 4y + 2 = 4(2x + 1) + 2 = 8x + 6.$$

The two computations of dz/dx therefore lead to the same answer. In this example there was no clear advantage in using the chain rule. The chain rule becomes useful when the functions f and g become more complicated.

² Briefly, you have to show that the function

$$h(y) = \begin{cases} \{f(y) - f(g(a))\}/(y - g(a)) & y \neq a \\ f'(g(a)) & y = a \end{cases}$$

is continuous.

13.4. Example where you really need the Chain Rule. We know what the derivative of $\sin x$ with respect to x is, but none of the rules we have found so far tell us how to differentiate $f(x) = \sin(2x)$.

The function $f(x) = \sin 2x$ is the composition of two simpler functions, namely

$$f(x) = g(h(x)) \text{ where } g(u) = \sin u \text{ and } h(x) = 2x.$$

We know how to differentiate each of the two functions g and h :

$$g'(u) = \cos u, \quad h'(x) = 2.$$

Therefore the chain rule implies that

$$f'(x) = g'(h(x))h'(x) = \cos(2x) \cdot 2 = 2 \cos 2x.$$

Leibniz would have decomposed the relation $y = \sin 2x$ between y and x as

$$y = \sin u, \quad u = 2x$$

and then computed the derivative of $\sin 2x$ with respect to x as follows

$$\frac{d \sin 2x}{dx} \stackrel{u=2x}{=} \frac{d \sin u}{dx} = \frac{d \sin u}{du} \cdot \frac{du}{dx} = \cos u \cdot 2 = 2 \cos 2x.$$

13.5. The Power Rule and the Chain Rule. The Power Rule, which says that for any function f and any rational number n one has

$$\frac{d}{dx}(f(x)^n) = n f(x)^{n-1} f'(x),$$

is a special case of the Chain Rule, for one can regard $y = f(x)^n$ as the composition of two functions

$$y = g(u), \quad u = f(x)$$

where $g(u) = u^n$. Since $g'(u) = nu^{n-1}$ the Chain Rule implies that

$$\frac{du^n}{dx} = \frac{du^n}{du} \cdot \frac{du}{dx} = nu^{n-1} \frac{du}{dx}.$$

Setting $u = f(x)$ and $\frac{du}{dx} = f'(x)$ then gives you the Power Rule.

13.6. The volume of an inflating balloon. Consider the “real world example” from page 53 again. There we considered a growing water balloon of radius

$$r = f(t).$$

The volume of this balloon is

$$V = \frac{4}{3}\pi r^3 = \frac{4}{3}\pi f(t)^3.$$

We can regard this as the composition of two functions, $V = g(r) = \frac{4}{3}\pi r^3$ and $r = f(t)$.

According to the chain rule the rate of change of the volume with time is now

$$\frac{dV}{dt} = \frac{dV}{dr} \frac{dr}{dt}$$

i.e. it is the product of the rate of change of the volume with the radius of the balloon and the rate of change of the balloon’s radius with time. From

$$\frac{dV}{dr} = \frac{d\frac{4}{3}\pi r^3}{dr} = 4\pi r^2$$

we see that

$$\frac{dV}{dr} = 4\pi r^2 \frac{dr}{dt}.$$

For instance, if the radius of the balloon is growing at 0.5inch/sec, and if its radius is $r = 3.0$ inch, then the volume is growing at a rate of

$$\frac{dV}{dt} = 4\pi(3.0\text{inch})^2 \times 0.5\text{inch/sec} \approx 57\text{inch}^3/\text{sec}.$$

13.7. A more complicated example. Suppose you needed to find the derivative of

$$y = h(x) = \frac{\sqrt{x+1}}{(\sqrt{x+1}+1)^2}$$

We can write this function as a composition of two simpler functions, namely,

$$y = f(u), \quad u = g(x),$$

with

$$f(u) = \frac{u}{(u+1)^2} \text{ and } g(x) = \sqrt{x+1}.$$

The derivatives of f and g are

$$f'(u) = \frac{1 \cdot (u+1)^2 - u \cdot 2(u+1)}{(u+1)^4} = \frac{u+1-2}{(u+1)^3} = \frac{u-1}{(u+1)^3},$$

and

$$g'(x) = \frac{1}{2\sqrt{x+1}}.$$

Hence the derivative of the composition is

$$h'(x) = \frac{d}{dx} \left\{ \frac{\sqrt{x+1}}{(\sqrt{x+1}+1)^2} \right\} = f'(u)g'(x) = \frac{u-1}{(u+1)^3} \cdot \frac{1}{2\sqrt{x+1}}.$$

The result should be a function of x , and we achieve this by replacing all u 's with $u = \sqrt{x+1}$:

$$\frac{d}{dx} \left\{ \frac{\sqrt{x+1}}{(\sqrt{x+1}+1)^2} \right\} = \frac{\sqrt{x+1}-1}{(\sqrt{x+1}+1)^3} \cdot \frac{1}{2\sqrt{x+1}}.$$

The last step (where you replace u by its definition in terms of x) is important because the problem was presented to you with only x and y as variables while u was a variable you introduced yourself to do the problem.

Sometimes it is possible to apply the Chain Rule without introducing new letters, and you will simply think “the derivative is the derivative of the outside with respect to the inside times the derivative of the inside.” For instance, to compute

$$\frac{d}{dx} \frac{4 + \sqrt{7+x^3}}{dx}$$

you could set $u = 7 + x^3$, and compute

$$\frac{d}{dx} \frac{4 + \sqrt{7+x^3}}{dx} = \frac{d}{du} \frac{4 + \sqrt{u}}{du} \cdot \frac{du}{dx}.$$

Instead of writing all this explicitly, you could think of $u = 7 + x^3$ as the function “inside the square root,” and think of $4 + \sqrt{u}$ as “the outside function.” You would then immediately write

$$\frac{d}{dx} (4 + \sqrt{7+x^3}) = \frac{1}{2\sqrt{7+x^3}} \cdot 3x^2.$$

13.8. The Chain Rule and composing more than two functions. Often we have to apply the Chain Rule more than once to compute a derivative. Thus if $y = f(u)$, $u = g(v)$, and $v = h(x)$ we have

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dv} \cdot \frac{dv}{dx}.$$

In functional notation this is

$$(f \circ g \circ h)'(x) = f'(g(h(x))) \cdot g'(h(x)) \cdot h'(x).$$

Note that each of the three derivatives on the right is evaluated at a different point. Thus if $b = h(a)$ and $c = g(b)$ the Chain Rule is

$$\left. \frac{dy}{dx} \right|_{x=a} = \left. \frac{dy}{du} \right|_{u=c} \cdot \left. \frac{du}{dv} \right|_{v=b} \cdot \left. \frac{dv}{dx} \right|_{x=a}.$$

For example, if $y = \frac{1}{1 + \sqrt{9 + x^2}}$, then $y = 1/(1 + u)$ where $u = 1 + \sqrt{v}$ and $v = 9 + x^2$ so

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dv} \cdot \frac{dv}{dx} = -\frac{1}{(1+u)^2} \cdot \frac{1}{2\sqrt{v}} \cdot 2x.$$

so

$$\left. \frac{dy}{dx} \right|_{x=4} = \left. \frac{dy}{du} \right|_{u=6} \cdot \left. \frac{du}{dv} \right|_{v=25} \cdot \left. \frac{dv}{dx} \right|_{x=4} = -\frac{1}{7} \cdot \frac{1}{10} \cdot 8.$$

14. Exercises

- 149.** Let $y = \sqrt{1 + x^3}$ and find dy/dx using the Chain Rule. Say what plays the role of $y = f(u)$ and $u = g(x)$.

- 150.** Repeat the previous exercise with

$$y = (1 + \sqrt{1 + x})^3.$$

- 151.** Alice and Bob differentiated $y = \sqrt{1 + x^3}$ with respect to x differently. Alice wrote $y = \sqrt{u}$ and $u = 1 + x^3$ while Bob wrote $y = \sqrt{1 + v}$ and $v = x^3$. Assuming neither one made a mistake, did they get the same answer?

- 152.** Let $y = u^3 + 1$ and $u = 3x + 7$. Find $\frac{dy}{dx}$ and $\frac{dy}{du}$. Express the former in terms of x and the latter in terms of u .

- 153.** Suppose that $f(x) = \sqrt{x}$, $g(x) = 1 + x^2$, $v(x) = f \circ g(x)$, $w(x) = g \circ f(x)$. Find formulas for $v(x)$, $w(x)$, $v'(x)$, and $w'(x)$.

Compute the following derivatives

154. $f(x) = \sin 2x - \cos 3x$

155. $f(x) = \sin \frac{\pi}{x}$

156. $f(x) = \sin(\cos 3x)$

157. $f(x) = \frac{\sin x^2}{x^2}$

158. $f(x) = \tan \sqrt{1 + x^2}$

159. $f(x) = \cos^2 x - \cos x^2$

160. Group Problem.

Moe is pouring water into a glass. At time t (seconds) the height of the water in the glass is $h(t)$ (inch). The ACME glass company, which made the glass, says that the volume in the glass to height h is $V = 1.2 h^2$ (fluid ounces).

(a) The water height in the glass is rising at 2 inch per second at the moment that the height is 2 inch. How fast is Moe pouring water into the glass?

(b) If Moe pours water at a rate of 1 ounce per second, then how fast is the water level in the glass going up when it is 3 inches?

(c) Moe pours water at 1 ounce per second, and at some moment the water level is going up at 0.5 inch per second. What is the water level at that moment?

- 161.** Find the derivative of $f(x) = x \cos \frac{\pi}{x}$ at the point C in Figure 3.

- 162.** Suppose that $f(x) = x^2 + 1$, $g(x) = x + 5$, and

$$v = f \circ g, \quad w = g \circ f, \quad p = f \cdot g, \quad q = g \cdot f.$$

Find $v(x)$, $w(x)$, $p(x)$, and $q(x)$.

163. Group Problem.

Suppose that the functions f and g and their derivatives with respect to x have the following values at $x = 0$ and $x = 1$.

| x | $f(x)$ | $g(x)$ | $f'(x)$ | $g'(x)$ |
|-----|--------|--------|---------|---------|
| 0 | 1 | 1 | 5 | 1/3 |
| 1 | 3 | -4 | -1/3 | -8/3 |

Define

$$v(x) = f(g(x)), \quad w(x) = g(f(x)),$$

$$p(x) = f(x)g(x), \quad q(x) = g(x)f(x).$$

Evaluate $v(0)$, $w(0)$, $p(0)$, $q(0)$, $v'(0)$ and $w'(0)$, $p'(0)$, $q'(0)$. If there is insufficient information to answer the question, so indicate.

- 164.** A differentiable function f satisfies $f(3) = 5$, $f(9) = 7$, $f'(3) = 11$ and $f'(9) = 13$. Find an equation for the tangent line to the curve $y = f(x^2)$ at the point $(x, y) = (3, 7)$.

- 165.** There is a function f whose second derivative satisfies

$$f''(x) = -64f(x).$$

(a) One such function is $f(x) = \sin ax$, provided you choose the right constant a : Which value should a have?

(b) For which choices of the constants A , a and b does the function $f(x) = A \sin(ax + b)$ satisfy (\dagger) ?

166. Group Problem.

A cubical sponge, hereafter referred to as 'Bob', is absorbing water, which causes him to expand. His side at time t is $S(t)$. His volume is $V(t)$.

(a) What is the relation between $S(t)$ and $V(t)$, i.e. can you find a function f so that $V(t) = f(S(t))$?

(b) Describe the meaning of the derivatives $S'(t)$ and $V'(t)$ in one plain english sentence each. If we measure lengths in inches and time in minutes, then what units do t , $S(t)$, $V(t)$, $S'(t)$ and $V'(t)$ have?

(c) What is the relation between $S'(t)$ and $V'(t)$?

(d) At the moment that Bob's volume is 8 cubic inches, he is absorbing water at a rate of 2 cubic inch per minute. How fast is his side $S(t)$ growing?

15. Implicit differentiation

15.1. The recipe. Recall that an implicitly defined function is a function $y = f(x)$ which is defined by an equation of the form

$$F(x, y) = 0.$$

We call this equation the **defining equation** for the function $y = f(x)$. To find $y = f(x)$ for a given value of x you must solve the defining equation $F(x, y) = 0$ for y .

Here is a recipe for computing the derivative of an implicitly defined function.

- (1) Differentiate the equation $F(x, y) = 0$; you may need the chain rule to deal with the occurrences of y in $F(x, y)$;
- (2) You can rearrange the terms in the result of step 1 so as to get an equation of the form

$$(30) \quad G(x, y) \frac{dy}{dx} + H(x, y) = 0,$$

where G and H are expressions containing x and y but not the derivative.

- (3) Solve the equation in step 2 for $\frac{dy}{dx}$:

$$(31) \quad \frac{dy}{dx} = -\frac{H(x, y)}{G(x, y)}$$

- (4) If you also have an explicit description of the function (i.e. a formula expressing $y = f(x)$ in terms of x) then you can substitute $y = f(x)$ in the expression (31) to get a formula for dy/dx in terms of x only.

Often no explicit formula for y is available and you can't take this last step. In that case (31) is as far as you can go.

Observe that by following this procedure you will get a formula for the derivative $\frac{dy}{dx}$ which contains both x and y .

15.2. Dealing with equations of the form $F_1(x, y) = F_2(x, y)$. If the implicit definition of the function is not of the form $F(x, y) = 0$ but rather of the form $F_1(x, y) = F_2(x, y)$ then you move all terms to the left hand side, and proceed as above. E.g. to deal with a function $y = f(x)$ which satisfies

$$y^2 + x = xy$$

you rewrite this equation as

$$y^2 + x - xy = 0$$

and set $F(x, y) = y^2 + x - xy$.

15.3. Example – Derivative of $\sqrt[4]{1-x^4}$. Consider the function

$$f(x) = \sqrt[4]{1-x^4}, \quad -1 \leq x \leq 1.$$

We will compute its derivative in two ways: first the direct method, and then using the method of implicit differentiation (i.e. the recipe above).

The direct approach goes like this:

$$\begin{aligned} f'(x) &= \frac{d(1-x^4)^{1/4}}{dx} \\ &= \frac{1}{4}(1-x^4)^{-3/4} \frac{d(1-x^4)}{dx} \\ &= \frac{1}{4}(1-x^4)^{-3/4} (-4x^3) \\ &= -\frac{x^3}{(1-x^4)^{3/4}} \end{aligned}$$

To find the derivative using implicit differentiation we must first find a nice implicit description of the function. For instance, we could decide to get rid of all roots or fractional exponents in the function and point out that $y = \sqrt[4]{1-x^4}$ satisfies the equation $y^4 = 1-x^4$. So our implicit description of the function $y = f(x) = \sqrt[4]{1-x^4}$ is

$$x^4 + y^4 - 1 = 0; \quad \text{The defining function is therefore } F(x, y) = x^4 + y^4 - 1$$

Differentiate both sides with respect to x (and remember that $y = f(x)$, so y here is a function of x), and you get

$$\frac{dx^4}{dx} + \frac{dy^4}{dx} - \frac{d1}{dx} = 0 \implies 4x^3 + 4y^3 \frac{dy}{dx} = 0.$$

The expressions G and H from equation (30) in the recipe are $G(x, y) = 4y^3$ and $H(x, y) = 4x^3$.

This last equation can be solved for dy/dx :

$$\frac{dy}{dx} = -\frac{x^3}{y^3}.$$

This is a nice and short form of the derivative, but it contains y as well as x . To express dy/dx in terms of x only, and remove the y dependency we use $y = \sqrt[4]{1-x^4}$. The result is

$$f'(x) = \frac{dy}{dx} = -\frac{x^3}{y^3} = -\frac{x^3}{(1-x^4)^{3/4}}.$$

15.4. Another example. Let f be a function defined by

$$y = f(x) \iff 2y + \sin y = x, \text{ i.e. } 2y + \sin y - x = 0.$$

For instance, if $x = 2\pi$ then $y = \pi$, i.e. $f(2\pi) = \pi$.

To find the derivative dy/dx we differentiate the defining equation

$$\frac{d(2y + \sin y - x)}{dx} = \frac{d0}{dx} \implies 2\frac{dy}{dx} + \cos y \frac{dy}{dx} - \frac{dx}{dx} = 0 \implies (2 + \cos y) \frac{dy}{dx} - 1 = 0.$$

Solve for $\frac{dy}{dx}$ and you get

$$f'(x) = \frac{1}{2 + \cos y} = \frac{1}{2 + \cos f(x)}.$$

If we were asked to find $f'(2\pi)$ then, since we know $f(2\pi) = \pi$, we could answer

$$f'(2\pi) = \frac{1}{2 + \cos \pi} = \frac{1}{2 - 1} = 1.$$

If we were asked $f'(\pi/2)$, then all we would be able to say is

$$f'(\pi/2) = \frac{1}{2 + \cos f(\pi/2)}.$$

To say more we would first have to find $y = f(\pi/2)$, which one does by solving

$$2y + \sin y = \frac{\pi}{2}.$$

15.5. Derivatives of Arc Sine and Arc Tangent. Recall that

$$y = \arcsin x \iff x = \sin y \text{ and } -\frac{\pi}{2} \leq y \leq \frac{\pi}{2},$$

and

$$y = \arctan x \iff x = \tan y \text{ and } -\frac{\pi}{2} < y < \frac{\pi}{2}.$$

15.6. Theorem.

$$\begin{aligned}\frac{d \arcsin x}{dx} &= \frac{1}{\sqrt{1-x^2}} \\ \frac{d \arctan x}{dx} &= \frac{1}{1+x^2}\end{aligned}$$

PROOF. If $y = \arcsin x$ then $x = \sin y$. Differentiate this relation

$$\frac{dx}{dx} = \frac{d \sin y}{dx}$$

and apply the chain rule. You get

$$1 = (\cos y) \frac{dy}{dx},$$

and hence

$$\frac{dy}{dx} = \frac{1}{\cos y}.$$

How do we get rid of the y on the right hand side? We know $x = \sin y$, and also $-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$. Therefore

$$\sin^2 y + \cos^2 y = 1 \implies \cos y = \pm \sqrt{1 - \sin^2 y} = \pm \sqrt{1 - x^2}.$$

Since $-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$ we know that $\cos y \geq 0$, so we must choose the positive square root. This leaves us with $\cos y = \sqrt{1 - x^2}$, and hence

$$\frac{dy}{dx} = \frac{1}{\sqrt{1-x^2}}.$$

The derivative of $\arctan x$ is found in the same way, and you should really do this yourself. \square

16. Exercises

For each of the following problems find the derivative $f'(x)$ if $y = f(x)$ satisfies the given equation. State what the expressions $F(x, y)$, $G(x, y)$ and $H(x, y)$ from the recipe in the beginning of this section are.

If you can find an explicit description of the function $y = f(x)$, say what it is.

167. $xy = \frac{\pi}{6}$

168. $\sin(xy) = \frac{1}{2}$

169. $\frac{xy}{x+y} = 1$

170. $x + y = xy$

171. $(y-1)^2 + x = 0$

172. $(y+1)^2 + y - x = 0$

173. $(y-x)^2 + x = 0$

174. $(y+x)^2 + 2y - x = 0$

175. $(y^2 - 1)^2 + x = 0$

176. $(y^2 + 1)^2 - x = 0$

177. $x^3 + xy + y^3 = 3$

178. $\sin x + \sin y = 1$

179. $\sin x + xy + y^5 = \pi$

180. $\tan x + \tan y = 1$

For each of the following explicitly defined functions find an implicit definition which does not involve taking roots. Then use this description to find the derivative dy/dx .

181. $y = f(x) = \sqrt{1-x}$

182. $y = f(x) = \sqrt[4]{x+x^2}$

183. $y = f(x) = \sqrt{1-\sqrt{x}}$

184. $y = f(x) = \sqrt[4]{x-\sqrt{x}}$

185. $y = f(x) = \sqrt[3]{\sqrt{2x+1} - x^2}$

186. $y = f(x) = \sqrt[4]{x+x^2}$

187. $y = f(x) = \sqrt[3]{x - \sqrt{2x + 1}}$

188. $y = f(x) = \sqrt[4]{\sqrt[3]{x}}$

189. Group Problem.

(Inverse trig review) Simplify the following expressions, and indicate for which values of x (or θ , or ...) your simplification is valid. In case of doubt, try plotting the function on a graphing calculator.

- | | |
|----------------------------|----------------------------|
| (a) $\sin \arcsin x$ | (e) $\tan \arctan z$ |
| (b) $\cos \arcsin x$ | (f) $\tan \arcsin \theta$ |
| (c) $\arctan(\tan \theta)$ | (g) $\arcsin(\sin \theta)$ |
| (d) $\cot \arctan x$ | (h) $\cot \arcsin x$ |

Now that you know the derivatives of \arcsin and \arctan , you can find the derivatives of the following functions. What are they?

190. $f(x) = \arcsin(2x)$

191. $f(x) = \arcsin \sqrt{x}$

192. $f(x) = \arctan(\sin x)$

193. $f(x) = \sin \arctan x$

194. $f(x) = (\arcsin x)^2$

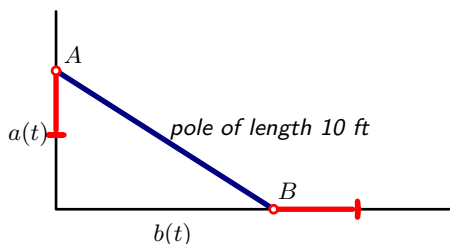
195. $f(x) = \frac{1}{1 + (\arctan x)^2}$

196. $f(x) = \sqrt{1 - (\arcsin x)^2}$

197. $f(x) = \frac{\arctan x}{\arcsin x}$

**PROBLEMS ON
RELATED RATES**

198. A 10 foot long pole has one end (B) on the floor and another (A) against a wall. If the bottom of the pole is 8 feet away from the wall, and if it is sliding away from the wall at 7 feet per second, then with what speed is the top (A) going down?



199. A pole 10 feet long rests against a vertical wall. If the bottom of the pole slides away from the wall at a speed of 2 ft/s, how fast is the angle between the top of the pole and the wall changing when the angle is $\pi/4$ radians?

200. A pole 13 meters long is leaning against a wall. The bottom of the pole is pulled along the ground away from the wall at the rate of 2 m/s. How fast is its height on the wall decreasing when the foot of the pole is 5 m away from the wall?

201. Group Problem.

A television camera is positioned 4000 ft from the base of a rocket launching pad. A rocket rises vertically and its speed is 600 ft/s when it has risen 3000 feet.

(a) How fast is the distance from the television camera to the rocket changing at that moment?

(b) How fast is the camera's angle of elevation changing at that same moment? (Assume that the television camera points toward the rocket.)

202. Group Problem.

A 2-foot tall dog is walking away from a streetlight which is on a 10-foot pole. At a certain moment, the tip of the dog's shadow is moving away from the streetlight at 5 feet per second. How fast is the dog walking at that moment?

203. An isosceles triangle is changing its shape: the lengths of the two equal sides remain fixed at 2 inch, but the angle $\theta(t)$ between them changes.

Let $A(t)$ be the area of the triangle at time t . If the area increases at a constant rate of $0.5 \text{ inch}^2/\text{sec}$, then how fast is the angle increasing or decreasing when $\theta = 60^\circ$?

204. A point P is moving in the first quadrant of the plane. Its motion is parallel to the x -axis; its distance to the x -axis is always 10 (feet). Its velocity is 3 feet per second to the left. We write θ for the angle between the positive x -axis and the line segment from the origin to P .

(a) Make a drawing of the point P .

(b) Where is the point when $\theta = \pi/3$?

(c) Compute the rate of change of the angle θ at the moment that $\theta = \frac{\pi}{3}$.

205. The point Q is moving on the line $y = x$ with velocity 3 m/sec. Find the rate of change of the following quantities at the moment in which Q is at the point $(1, 1)$:

(a) the distance from Q to the origin,

(b) the distance from Q to the point $R(2, 0)$,

(c) the angle $\angle ORQ$ where R is again the point $R(2, 0)$.

206. A point P is sliding on the parabola with equation $y = x^2$. Its x -coordinate is increasing at a constant rate of 2 feet/minute.

Find the rate of change of the following quantities at the moment that P is at $(3, 9)$:

(a) the distance from P to the origin,

(b) the area of the rectangle whose lower left corner is the origin and whose upper right corner is P ,

- (c) the slope of the tangent to the parabola at P ,
- (d) the angle $\angle OPQ$ where Q is the point $(0, 3)$.

207. Group Problem.

A certain amount of gas is trapped in a cylinder with a piston. The **ideal gas law** from thermodynamics says that if the cylinder is not heated, and if the piston moves slowly, then one has

$$pV = CT$$

where p is the pressure in the gas, V is its volume, T its temperature (in degrees Kelvin) and C is a constant depending on the amount of gas trapped in the cylinder.

(a) If the pressure is 10psi (pounds per square inch), if the volume is 25inch^3 , and if the piston is moving so

that the gas volume is expanding at a rate of 2inch^3 per minute, then what is the rate of change of the pressure?

(b) The ideal gas law turns out to be only approximately true. A more accurate description of gases is given by **van der Waals' equation of state**, which says that

$$\left(p + \frac{a}{V^2}\right)(V - b) = C$$

where a, b, C are constants depending on the temperature and the amount and type of gas in the cylinder.

Suppose that the cylinder contains fictitious gas for which one has $a = 12$ and $b = 3$. Suppose that at some moment the volume of gas is 12in^3 , the pressure is 25psi and suppose the gas is expanding at 2inch^3 per minute. Then how fast is the pressure changing?

Graph Sketching and Max-Min Problems

The signs of the first and second derivatives of a function tell us something about the shape of its graph. In this chapter we learn how to find that information.

1. Tangent and Normal lines to a graph

The slope of the tangent to the graph of f at the point $(a, f(a))$ is

$$(32) \quad m = f'(a)$$

and hence the equation for the tangent is

$$(33) \quad y = f(a) + f'(a)(x - a).$$

The slope of the normal line to the graph is $-1/m$ and thus one could write the equation for the normal as

$$(34) \quad y = f(a) - \frac{x - a}{f'(a)}.$$

When $f'(a) = 0$ the tangent is horizontal, and hence the normal is vertical. In this case the equation for the normal cannot be written as in (34), but instead one gets the simpler equation

$$y = f(a).$$

Both cases are covered by this form of the equation for the normal

$$(35) \quad x = a + f'(a)(f(a) - y).$$

Both (35) and (34) are formulas that you shouldn't try to remember. It is easier to remember that if the slope of the tangent is $m = f'(a)$, then the slope of the normal is $-1/m$.

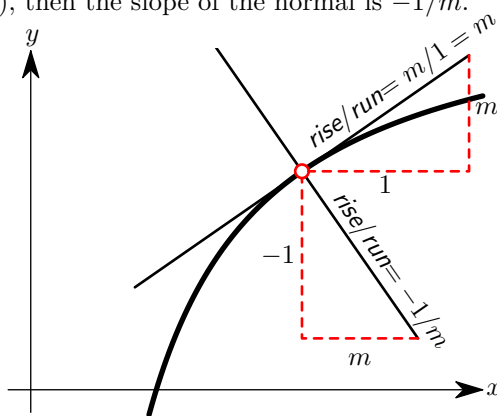


Figure 1. Why “slope of normal = $\frac{-1}{\text{slope of tangent}}$.”

2. The Intermediate Value Theorem

It is said that a function is continuous if you can draw its graph without taking your pencil off the paper. A more precise version of this statement is the *Intermediate Value Theorem*:

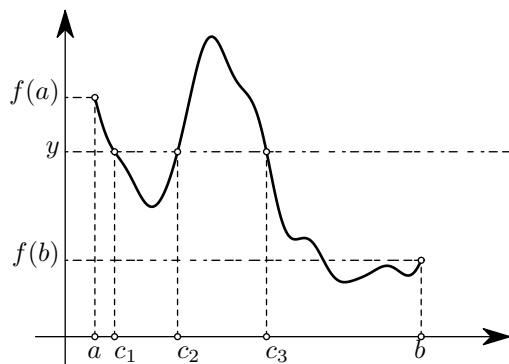


Figure 2. The Intermediate Value Theorem says that a continuous function must attain any given value y between $f(a)$ and $f(b)$ at least once. In this example there are three values of c for which $f(c) = y$ holds.

2.1. Intermediate Value Theorem. If f is a continuous function on an interval $a \leq x \leq b$, and if y is some number between $f(a)$ and $f(b)$, then there is a number c with $a \leq c \leq b$ such that $f(c) = y$.

Here “ y between $f(a)$ and $f(b)$ ” means that $f(a) \leq y \leq f(b)$ if $f(a) \leq f(b)$, and $f(b) \leq y \leq f(a)$ if $f(b) \leq f(a)$.

2.2. Example – Square root of 2. Consider the function $f(x) = x^2$. Since $f(1) < 2$ and $f(2) = 4 > 2$ the intermediate value theorem with $a = 1$, $b = 2$, $y = 2$ tells us that there is a number c between 1 and 2 such that $f(c) = 2$, i.e. for which $c^2 = 2$. So the theorem tells us that the square root of 2 exists.

2.3. Example – The equation $\theta + \sin \theta = \frac{\pi}{2}$. Consider the function $f(x) = x + \sin x$. It is a continuous function at all x , so from $f(0) = 0$ and $f(\pi) = \pi$ it follows that there is a number θ between 0 and π such that $f(\theta) = \pi/2$. In other words, the equation

$$(36) \quad \theta + \sin \theta = \frac{\pi}{2}$$

has a solution θ with $0 \leq \theta \leq \pi/2$. Unlike the previous example, where we knew the solution was $\sqrt{2}$, there is no simple formula for the solution to (36).

2.4. Example – Solving $1/x = 0$. If we apply the intermediate value theorem to the function $f(x) = 1/x$ on the interval $[a, b] = [-1, 1]$, then we see that for any y between $f(a) = f(-1) = -1$ and $f(b) = f(1) = 1$ there is a number c in the interval $[-1, 1]$ such that $1/c = y$. For instance, we could choose $y = 0$ (that’s between -1 and $+1$), and conclude that there is some c with $-1 \leq c \leq 1$ and $1/c = 0$.

But there is no such c , because $1/c$ is never zero! So we have done something wrong, and the mistake we made is that we overlooked that our function $f(x) = 1/x$ is not defined on the **whole** interval $-1 \leq x \leq 1$ because it is not defined at $x = 0$. *The moral: always check the hypotheses of a theorem before you use it!*

3. Exercises

208. Where does the normal to the graph of $y = x^2$ at the point $(1, 1)$ intersect the x -axis?
209. Where does the tangent to the graph of $y = x^2$ at the point (a, a^2) intersect the x -axis?
210. Where does the normal to the graph of $y = x^2$ at the point (a, a^2) intersect the x -axis?
211. Where does the normal to the graph of $y = \sqrt{x}$ at the point (a, \sqrt{a}) intersect the x -axis?
212. Does the graph of $y = x^4 - 2x^2 + 2$ have any horizontal tangents? If so, where?
Does the graph of the same function have any vertical tangents?
Does it have vertical normals?
Does it have horizontal normals?
213. At some point $(a, f(a))$ on the graph of $f(x) = -1 + 2x - x^2$ the tangent to this graph goes through the origin. Which point is it?

214. Find equations for the tangent and normal lines

to the curve ... at the point...

- | | |
|------------------------|--------|
| (a) $y = 4x/(1 + x^2)$ | (1, 2) |
| (b) $y = 8/(4 + x^2)$ | (2, 1) |
| (c) $y^2 = 2x + x^2$ | (2, 2) |
| (d) $xy = 3$ | (1, 3) |

satisfies $f(-1) = -2$ and $f(+1) = +2$, so, by the Intermediate Value Theorem, there should be some value c between -1 and $+1$ such that $f(c) = 0$. **True or False?**

216. Find the equation for the tangents to the graph of the Backward Sine at the points $x = 1$, $x = \frac{1}{2}$ and at D (see Figure 2 in §7.3.)

215. Group Problem.

The function

$$f(x) = \frac{x^2 + |x|}{x}$$

217. Find the equation for the tangent to the graph of the Backward Cosine in a Bow Tie at the point C (see Figure 3 in §9.3)

4. Finding sign changes of a function

The intermediate value theorem implies the following very useful fact.

4.1. Theorem. *If f is continuous function on some interval $a < x < b$, and if $f(x) \neq 0$ for all x in this interval, then $f(x)$ is either positive for all $a < x < b$ or else it is negative for all $a < x < b$.*

PROOF. The theorem says that there can't be two numbers $a < x_1 < x_2 < b$ such that $f(x_1)$ and $f(x_2)$ have opposite signs. If there were two such numbers then the intermediate value theorem would imply that somewhere between x_1 and x_2 there was a c with $f(c) = 0$. But we are assuming that $f(c) \neq 0$ whenever $a < c < b$. \square

4.2. Example. Consider

$$f(x) = (x - 3)(x - 1)^2(2x + 1)^3.$$

The zeros of f (i.e. the solutions of $f(x) = 0$) are $-\frac{1}{2}, 1, 3$. These numbers split the real line into four intervals

$$(-\infty, -\frac{1}{2}), \quad (-\frac{1}{2}, 1), \quad (1, 3), \quad (3, \infty).$$

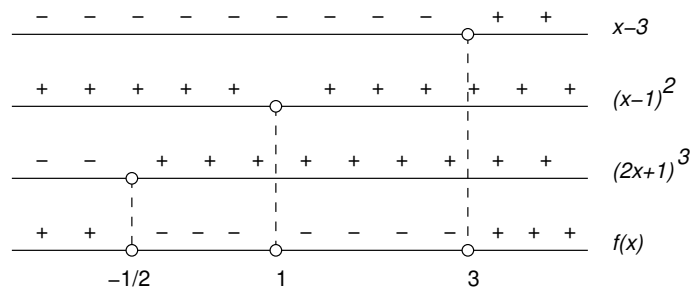
Theorem 4.1 tells us that $f(x)$ cannot change its sign in any of these intervals. For instance, $f(x)$ has the same sign for all x in the first interval $(-\infty, -\frac{1}{2})$. Now we choose a number we like from this interval (e.g. -1) and find the sign of $f(-1)$: $f(-1) = (-4)(-2)^2(-3)^3$ is positive. Therefore $f(x) > 0$ for all x in the interval $(-\infty, -\frac{1}{2})$. In the same we find

$$\begin{aligned} f(-1) &= (-4)(-2)^2(-3)^3 > 0 &\implies f(x) > 0 \text{ for } x < -\frac{1}{2} \\ f(0) &= (-3)(-1)^2(1)^3 < 0 &\implies f(x) < 0 \text{ for } -\frac{1}{2} < x < 1 \\ f(2) &= (-1)(1)^2(5)^3 < 0 &\implies f(x) < 0 \text{ for } 1 < x < 3 \\ f(4) &= (1)(3)^2(9)^3 > 0 &\implies f(x) > 0 \text{ for } x > 3. \end{aligned}$$

If you know all the zeroes of a continuous function, then this method allows you to decide where the function is positive or negative. However, when the given function is factored into easy functions, as in this example, there is a different way of finding the signs of f . For each of the factors $x - 3$, $(x - 1)^2$ and $(2x + 1)^3$ it is easy to determine the sign, for any given x . These signs can only change at a zero of the factor. Thus we have

- $x - 3$ is positive for $x > 3$ and negative for $x < 3$;
- $(x - 1)^2$ is always positive (except at $x = 1$);
- $(2x + 1)^3$ is positive for $x > -\frac{1}{2}$ and negative for $x < -\frac{1}{2}$.

Multiplying these signs we get the same conclusions as above. We can summarize this computation in the following diagram:



5. Increasing and decreasing functions

Here are four very similar definitions – look closely to see how they differ.

- A function is called **increasing** if $a < b$ implies $f(a) < f(b)$ for all numbers a and b in the domain of f .
- A function is called **decreasing** if $a < b$ implies $f(a) > f(b)$ for all numbers a and b in the domain of f .
- The function f is called **non-decreasing** if $a < b$ implies $f(a) \leq f(b)$ for all numbers a and b in the domain of f .
- The function f is called **non-increasing** if $a < b$ implies $f(a) \geq f(b)$ for all numbers a and b in the domain of f .

You can summarize these definitions as follows:

| f is ... | if for all a and b one has... |
|-----------------|-----------------------------------|
| Increasing: | $a < b \implies f(a) < f(b)$ |
| Decreasing: | $a < b \implies f(a) > f(b)$ |
| Non-increasing: | $a < b \implies f(a) \geq f(b)$ |
| Non-decreasing: | $a < b \implies f(a) \leq f(b)$ |

The sign of the derivative of f tells you if f is increasing or not. More precisely:

5.1. Theorem. *If a function is non-decreasing on an interval $a < x < b$ then $f'(x) \geq 0$ for all x in that interval.*

If a function is non-increasing on an interval $a < x < b$ then $f'(x) \leq 0$ for all x in that interval.

For instance, if f is non-decreasing, then for any given x and any positive Δx one has $f(x + \Delta x) \geq f(x)$ and hence

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} \geq 0.$$

Now let $\Delta x \searrow 0$ and you find that

$$f'(x) = \lim_{\Delta x \searrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \geq 0.$$

What about the converse, i.e. if you know the sign of f' then what can you say about f ? For this we have the following

5.2. Theorem. *Suppose f is a differentiable function on an interval (a, b) .*

If $f'(x) > 0$ for all $a < x < b$, then f is increasing.

If $f'(x) < 0$ for all $a < x < b$, then f is decreasing.

The proof is based on the Mean Value theorem which also finds use in many other situations:

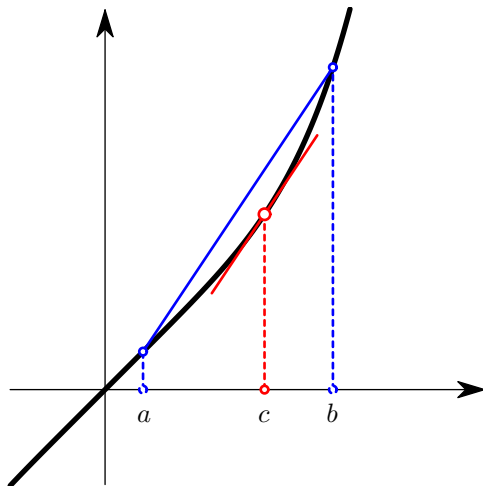


Figure 3. According to the Mean Value Theorem there always is some number c between a and b such that the tangent to the graph of f is parallel to the line segment connecting the two points $(a, f(a))$ and $(b, f(b))$. This is true for any choice of a and b ; c depends on a and b of course.

5.3. The Mean Value Theorem. *If f is a differentiable function on the interval $a \leq x \leq b$, then there is some number c , with $a < c < b$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

PROOF OF THEOREM 5.2. We show that $f'(x) > 0$ for all x implies that f is increasing. Let $x_1 < x_2$ be two numbers between a and b . Then the Mean Value Theorem implies that there is some c between x_1 and x_2 such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1},$$

or

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Since we know that $f'(c) > 0$ and $x_2 - x_1 > 0$ it follows that $f(x_2) - f(x_1) > 0$, i.e. $f(x_2) > f(x_1)$. □

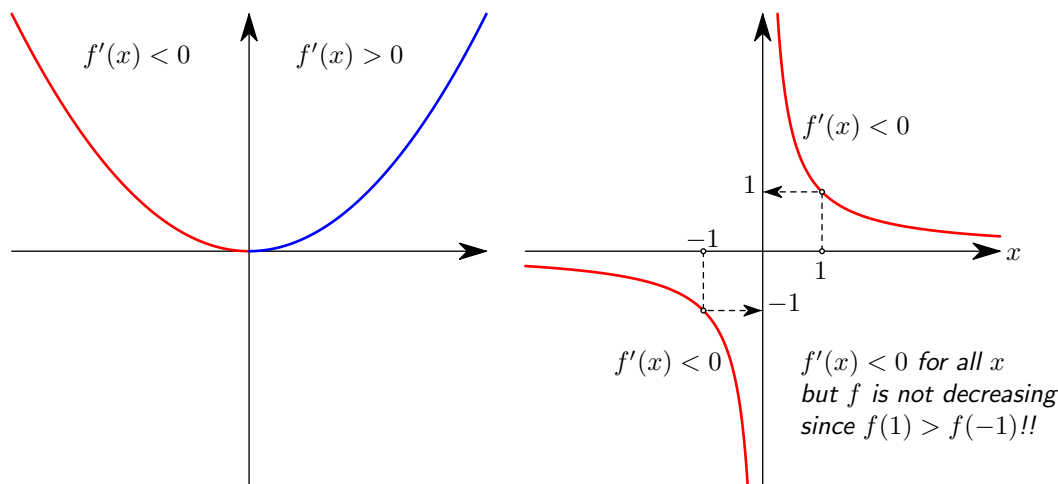
6. Examples

Armed with these theorems we can now split the graph of any function into increasing and decreasing parts simply by computing the derivative $f'(x)$ and finding out where $f'(x) > 0$ and where $f'(x) < 0$ – i.e. we apply the method from the previous section to f' rather than f .

6.1. Example: the parabola $y = x^2$. The familiar graph of $f(x) = x^2$ consists of two parts, one decreasing and one increasing. You can see this from the derivative which is

$$f'(x) = 2x \begin{cases} > 0 & \text{for } x > 0 \\ < 0 & \text{for } x < 0. \end{cases}$$

Therefore the function $f(x) = x^2$ is **decreasing** for $x < 0$ and **increasing** for $x > 0$.



6.2. Example: the hyperbola $y = 1/x$. The derivative of the function $f(x) = 1/x = x^{-1}$ is

$$f'(x) = -\frac{1}{x^2}$$

which is always negative. You would therefore think that this function is decreasing, or at least non-increasing: if $a < b$ then $1/a \geq 1/b$. But this isn't true if you take $a = -1$ and $b = 1$:

$$a = -1 < 1 = b, \text{ but } \frac{1}{a} = -1 < 1 = \frac{1}{b} !!$$

The problem is that we used theorem 5.2, but if you carefully read that theorem then you see that it applies to functions **that are defined on an interval**. The function in this example, $f(x) = 1/x$, is not defined on the interval $-1 < x < 1$ because it isn't defined at $x = 0$. That's why you can't conclude that the $f(x) = 1/x$ is increasing from $x = -1$ to $x = +1$.

On the other hand, the function is defined and differentiable on the interval $0 < x < \infty$, so theorem 5.2 tells us that $f(x) = 1/x$ is decreasing for $x > 0$. This means, that as long as x is positive, increasing x will decrease $1/x$.

6.3. Graph of a cubic function. Consider the function

$$y = f(x) = x^3 - x.$$

Its derivative is

$$f'(x) = 3x^2 - 1.$$

We try to find out where f' is positive, and where it is negative by factoring $f'(x)$

$$f'(x) = 3\left(x^2 - \frac{1}{3}\right) = 3\left(x + \frac{1}{3}\sqrt{3}\right)\left(x - \frac{1}{3}\sqrt{3}\right)$$

from which you see that

$$f'(x) > 0 \text{ for } x < -\frac{1}{3}\sqrt{3}$$

$$f'(x) < 0 \text{ for } -\frac{1}{3}\sqrt{3} < x < \frac{1}{3}\sqrt{3}$$

$$f'(x) > 0 \text{ for } x > \frac{1}{3}\sqrt{3}$$

Therefore the function f is

$$\text{increasing on } \left(-\infty, -\frac{1}{3}\sqrt{3}\right), \quad \text{decreasing on } \left(-\frac{1}{3}\sqrt{3}, \frac{1}{3}\sqrt{3}\right), \quad \text{increasing on } \left(\frac{1}{3}\sqrt{3}, \infty\right).$$

At the two points $x = \pm\frac{1}{3}\sqrt{3}$ one has $f'(x) = 0$ so there the tangent will be horizontal. This leads us to the following picture of the graph of f :

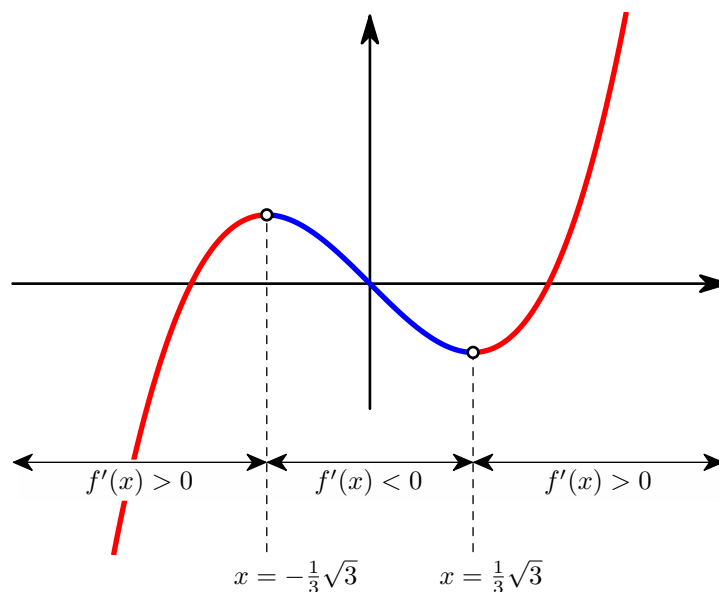


Figure 4. The graph of $f(x) = x^3 - x$.

6.4. A function whose tangent turns up and down infinitely often near the origin. We end with a weird example. Somewhere in the mathematician's zoo of curious functions the following will be on exhibit. Consider the function

$$f(x) = \frac{x}{2} + x^2 \sin \frac{\pi}{x}.$$

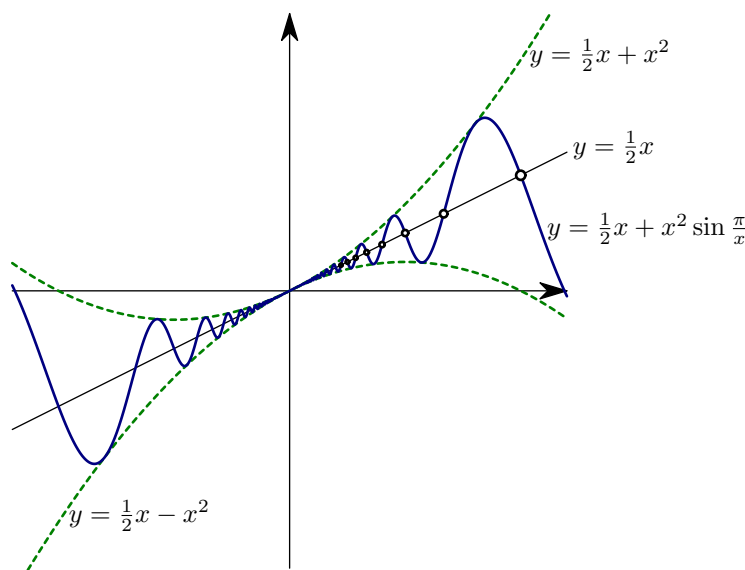


Figure 5. Positive derivative at a point ($x = 0$) does not mean that the function is “increasing near that point.” The slopes at the intersection points alternate between $\frac{1}{2} - \pi$ and $\frac{1}{2} + \pi$.

For $x = 0$ this formula is undefined, and we are free to define $f(0) = 0$. This makes the function continuous at $x = 0$. In fact, this function is differentiable at $x = 0$, with derivative given by

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0} \frac{1}{2} + x \sin \frac{\pi}{x} = \frac{1}{2}.$$

(To find the limit apply the sandwich theorem to $-|x| \leq x \sin \frac{\pi}{x} \leq |x|$.)

So the slope of the tangent to the graph at the origin is positive ($\frac{1}{2}$), and one would *think* that the function should be increasing near $x = 0$ (i.e. bigger x gives bigger $f(x)$.) The point of this example is that this turns out not to be true.

To explain why not, we must compute the derivative of this function for $x \neq 0$. It is given by

$$f'(x) = \frac{1}{2} - \pi \cos \frac{\pi}{x} + 2x \sin \frac{\pi}{x}.$$

Now consider the sequence of intersection points P_1, P_2, \dots of the graph with the line $y = x/2$. They are

$$P_k(x_k, y_k), \quad x_k = \frac{1}{k}, \quad y_k = f(x_k).$$

For larger and larger k the points P_k tend to the origin (the x coordinate is $\frac{1}{k}$ which goes to 0 as $k \rightarrow \infty$). The slope of the tangent at P_k is given by

$$\begin{aligned} f'(x_k) &= \frac{1}{2} - \pi \cos \frac{\pi}{1/k} + 2 \frac{1}{k} \sin \frac{\pi}{1/k} \\ &= \frac{1}{2} - \pi \underbrace{\cos k\pi}_{=(-1)^k} + \frac{2}{k} \underbrace{\sin k\pi}_{=0} \\ &= \begin{cases} -\frac{1}{2} - \pi \approx -2.64159265358979\dots & \text{for } k \text{ even} \\ \frac{1}{2} + \pi \approx +3.64159265358979\dots & \text{for } k \text{ odd} \end{cases} \end{aligned}$$

In other words, along the sequence of points P_k the slope of the tangent flip-flops between $\frac{1}{2} - \pi$ and $\frac{1}{2} + \pi$, i.e. between a positive and a negative number.

In particular, the slope of the tangent at the odd intersection points is negative, and so you would expect the function to be decreasing there. In other words we see that ***even though the derivative at $x = 0$ of this function is positive, there are points on the graph arbitrarily close to the origin where the tangent has negative slope.***

7. Maxima and Minima

A function has a **global maximum** at some a in its domain if $f(x) \leq f(a)$ for all other x in the domain of f . Global maxima are sometimes also called “absolute maxima.”

A function has a **local maximum** at some a in its domain if there is a small $\delta > 0$ such that $f(x) \leq f(a)$ for all x with $a - \delta < x < a + \delta$ which lie in the domain of f .

Every global maximum is a local maximum, but a local maximum doesn't have to be a global maximum.

7.1. Where to find local maxima and minima. Any x value for which $f'(x) = 0$ is called a **stationary point** for the function f .

7.2. Theorem. Suppose f is a differentiable function on some interval $[a, b]$.

Every local maximum or minimum of f is either one of the end points of the interval $[a, b]$, or else it is a stationary point for the function f .

PROOF. Suppose that f has a local maximum at x and suppose that x is not a or b . By assumption the left and right hand limits

$$f'(x) = \lim_{\Delta x \nearrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \quad \text{and} \quad f'(x) = \lim_{\Delta x \searrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

both exist and they are equal.

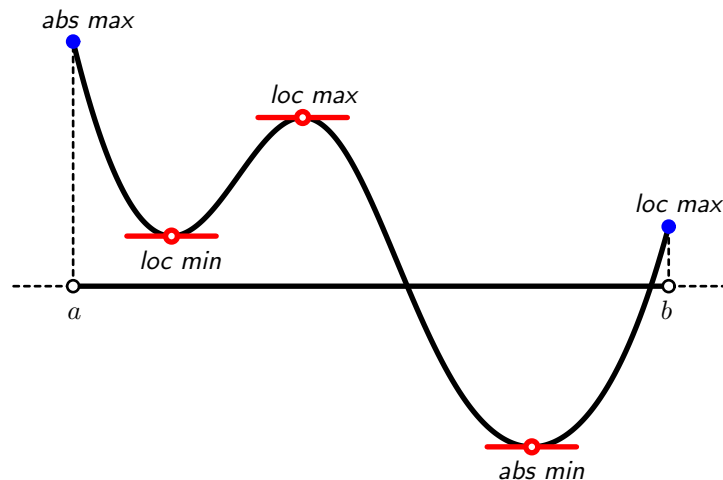


Figure 6. A function defined on an interval $[a, b]$ with one interior absolute minimum, another interior local minimum, an interior local maximum, and two local maxima on the boundary, one of which is in fact an absolute maximum.

Since f has a local maximum at x we have $f(x + \Delta x) - f(x) \leq 0$ if $-\delta < \Delta x < \delta$. In the first limit we also have $\Delta x < 0$, so that

$$\lim_{\Delta x \nearrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \leq 0$$

Hence $f'(x) \leq 0$.

In the second limit we have $\Delta x > 0$, so

$$\lim_{\Delta x \searrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \geq 0$$

which implies $f'(x) \geq 0$.

Thus we have shown that $f'(x) \leq 0$ and $f'(x) \geq 0$ at the same time. This can only be true if $f'(x) = 0$. \square

7.3. How to tell if a stationary point is a maximum, a minimum, or neither. If $f'(c) = 0$ then c is a stationary point (by definition), and it might be local maximum or a local minimum. You can tell what kind of stationary point c is by looking at the signs of $f'(x)$ for x near c .

7.4. Theorem. If in some small interval $(c - \delta, c + \delta)$ you have $f'(x) < 0$ for $x < c$ and $f'(x) > 0$ for $x > c$ then f has a local minimum at $x = c$.

If in some small interval $(c - \delta, c + \delta)$ you have $f'(x) > 0$ for $x < c$ and $f'(x) < 0$ for $x > c$ then f has a local maximum at $x = c$. The reason is simple: if f increases to the left of c and decreases to the right of c then it has a maximum at c . More precisely:

if $f'(x) > 0$ for x between $c - \delta$ and c , then f is increasing for $c - \delta < x < c$ and therefore $f(x) < f(c)$ for x between $c - \delta$ and c .

If in addition $f'(x) < 0$ for $x > c$ then f is decreasing for x between c and $c + \delta$, so that $f(x) < f(c)$ for those x .

Combine these two facts and you get $f(x) \leq f(c)$ for $c - \delta < x < c + \delta$.

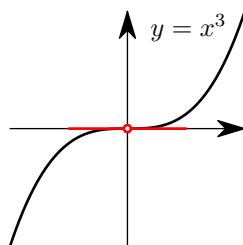
7.5. Example – local maxima and minima of $f(x) = x^3 - x$. In §6.3 we had found that the function $f(x) = x^3 - x$ is decreasing when $-\infty < x < -\frac{1}{3}\sqrt{3}$, and also when $\frac{1}{3}\sqrt{3} < x < \infty$, while it is increasing when $-\frac{1}{3}\sqrt{3} < x < \frac{1}{3}\sqrt{3}$. It follows that the function has a local minimum at $x = -\frac{1}{3}\sqrt{3}$, and a local maximum at $x = \frac{1}{3}\sqrt{3}$.

Neither the local maximum nor the local minimum are global max or min since

$$\lim_{x \rightarrow -\infty} f(x) = +\infty \text{ and } \lim_{x \rightarrow \infty} f(x) = -\infty.$$

7.6. A stationary point that is neither a maximum nor a minimum. If you look for stationary points of the function $f(x) = x^3$ you find that there's only one, namely $x = 0$. The derivative $f'(x) = 3x^2$ does not change sign at $x = 0$, so the test in Theorem 7.4 does not tell us anything.

And in fact, $x = 0$ is neither a local maximum nor a local minimum since $f(x) < f(0)$ for $x < 0$ and $f(x) > 0$ for $x > 0$.



8. Must there always be a maximum?

Theorem 7.2 is very useful since it tells you how to find (local) maxima and minima. The following theorem is also useful, but in a different way. It doesn't say how to find maxima or minima, but it tells you that they do exist, and hence that you are not wasting your time trying to find a maximum or minimum.

8.1. Theorem. *Let f be continuous function defined on the **closed** interval $a \leq x \leq b$. Then f attains its maximum and also its minimum somewhere in this interval. In other words there exist real numbers c and d such that*

$$f(c) \leq f(x) \leq f(d)$$

whenever $a \leq x \leq b$. The proof of this theorem requires a more careful definition of the real numbers than we have given in Chapter 1, and we will take the theorem for granted.

9. Examples – functions with and without maxima or minima

In the following three example we explore what can happen if some of the hypotheses in Theorem 8.1 are not met.

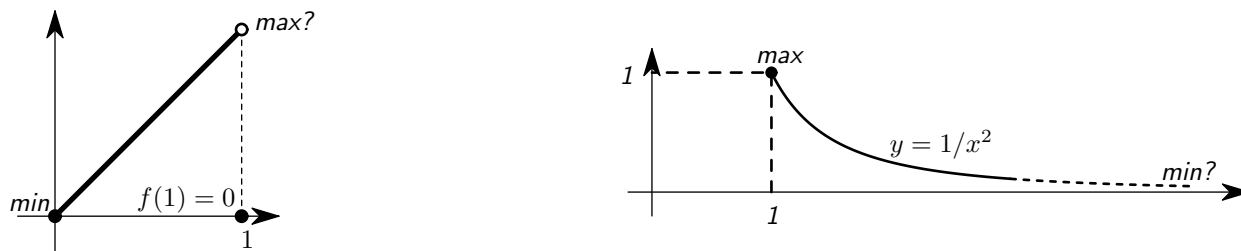


Figure 7. The function on the left has no maximum, and the one on the right has no minimum.

9.1. Question: Does the function

$$f(x) = \begin{cases} x & \text{for } 0 \leq x < 1 \\ 0 & \text{for } x = 1. \end{cases}$$

have a maximum on the interval $0 \leq x \leq 1$?

Answer: No. What would the maximal value be? Since

$$\lim_{x \nearrow 1} f(x) = \lim_{x \nearrow 1} x = 1$$

The maximal value cannot be less than 1. On the other hand the function is never larger than 1. So if there were a number a in the interval $[0, 1]$ such that $f(a)$ was the maximal value of f , then we would have $f(a) = 1$. If you now search the interval for numbers a with $f(a) = 1$, then you notice that such an a does not exist. Conclusion: this function does **not** attain its maximum on the interval $[0, 1]$.

What about Theorem 8.1? That theorem only applies to continuous functions, and the function f in this example is not continuous at $x = 1$. For at $x = 1$ one has

$$f(1) = 0 \neq 1 = \lim_{x \nearrow 1} f(x).$$

So all it takes for the Theorem to fail is that the function f be discontinuous at just one point.

9.2. Question: Does the function

$$f(x) = \frac{1}{x^2}, \quad 1 \leq x < \infty$$

have a maximum or minimum?

Answer: The function has a maximum at $x = 1$, but it has no minimum.

Concerning the maximum: if $x \geq 1$ then $f(x) = 1/x^2 \leq 1$, while $f(1) = 1$. Hence $f(x) \leq f(1)$ for all x in the interval $[1, \infty)$ and that is why f attains its maximum at $x = 1$.

If we look for a minimal value of f then we note that $f(x) \geq 0$ for all x in the interval $[1, \infty)$, and also that

$$\lim_{x \rightarrow \infty} f(x) = 0,$$

so that **if** f attains a minimum at some a with $1 \leq a < \infty$, then the minimal value $f(a)$ must be zero. However, the equation $f(a) = 0$ has no solution – f does not attain its minimum.

Why does Theorem 8.1 not apply? In this example the function f is continuous on the whole interval $[1, \infty)$, but this interval is not a closed interval, i.e. it is not of the form $[a, b]$ (it does not include its endpoints).

10. General method for sketching the graph of a function

Given a differentiable function f defined on some interval $a \leq x \leq b$, you can find the increasing and decreasing parts of the graph, as well as all the local maxima and minima by following this procedure:

- (1) find all solutions of $f'(x) = 0$ in the interval $[a, b]$: these are called the *critical* or *stationary* points for f .
- (2) find the sign of $f'(x)$ at all other points
- (3) each stationary point at which $f'(x)$ actually changes sign is a local maximum or local minimum. Compute the function value $f(x)$ at each stationary point.
- (4) compute the function values at the endpoints of the interval, i.e. compute $f(a)$ and $f(b)$.
- (5) the absolute maximum is attained at the stationary point or the boundary point with the highest function value; the absolute minimum occurs at the boundary or stationary point with the smallest function value.

If the interval is unbounded, i.e. if the function is defined for $-\infty < x < \infty$ then you can't compute the values $f(a)$ and $f(b)$, but instead you should compute $\lim_{x \rightarrow \pm\infty} f(x)$.

10.1. Example – the graph of a rational function. Let's “sketch the graph” of the function

$$f(x) = \frac{x(1-x)}{1+x^2}.$$

By looking at the signs of numerator and denominator we see that

$$\begin{aligned} f(x) &> 0 \text{ for } 0 < x < 1 \\ f(x) &< 0 \text{ for } x < 0 \text{ and also for } x > 1. \end{aligned}$$

We compute the derivative of f

$$f'(x) = \frac{1-2x-x^2}{(1+x^2)^2}.$$

Hence $f'(x) = 0$ holds if and only if

$$1-2x-x^2=0$$

and the solutions to this quadratic equation are $-1 \pm \sqrt{2}$. These two roots will appear several times and it will shorten our formulas if we abbreviate

$$A = -1 - \sqrt{2} \text{ and } B = -1 + \sqrt{2}.$$

To see if the derivative changes sign we factor the numerator and denominator. The denominator is always positive, and the numerator is

$$-x^2 - 2x + 1 = -(x^2 + 2x - 1) = -(x - A)(x - B).$$

Therefore

$$f'(x) \begin{cases} < 0 & \text{for } x < A \\ > 0 & \text{for } A < x < B \\ < 0 & \text{for } x > B \end{cases}$$

It follows that f is decreasing on the interval $(-\infty, A)$, increasing on the interval (A, B) and decreasing again on the interval (B, ∞) . Therefore

A is a local minimum, and B is a local maximum.

Are these global maxima and minima?

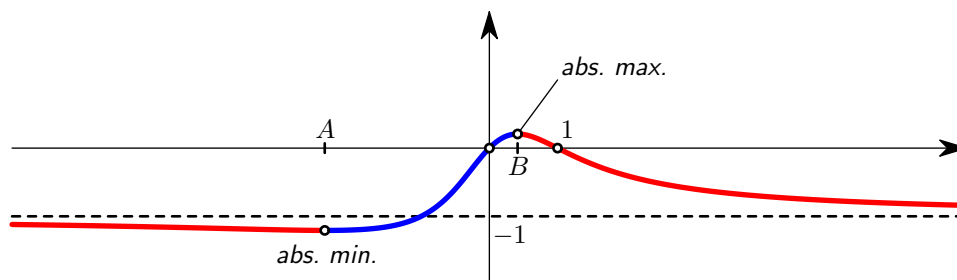


Figure 8. The graph of $f(x) = (x - x^2)/(1 + x^2)$

Since we are dealing with an unbounded interval we must compute the limits of $f(x)$ as $x \rightarrow \pm\infty$. You find

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = -1.$$

Since f is decreasing between $-\infty$ and A , it follows that

$$f(A) \leq f(x) < -1 \text{ for } -\infty < x \leq A.$$

Similarly, f is decreasing from B to $+\infty$, so

$$-1 < f(x) \leq f(-1 + \sqrt{2}) \text{ for } B < x < \infty.$$



Figure 9. If a graph is convex then all chords lie above the graph. If it is not convex then some chords will cross the graph or lie below it.

Between the two stationary points the function is increasing, so

$$f(-1 - \sqrt{2}) \leq f(x) \leq f(B) \text{ for } A \leq x \leq B.$$

From this it follows that $f(x)$ is the smallest it can be when $x = A = -1 - \sqrt{2}$ and at its largest when $x = B = -1 + \sqrt{2}$: the local maximum and minimum which we found are in fact a global maximum and minimum.

11. Convexity, Concavity and the Second Derivative

By definition, a function f is **convex** on some interval $a < x < b$ if the line segment connecting any pair of points on the graph lies *above* the piece of the graph between those two points.

The function is called **concave** if the line segment connecting any pair of points on the graph lies *below* the piece of the graph between those two points.

A point on the graph of f where $f''(x)$ changes sign is called an **inflection point**.

Instead of “convex” and “concave” one often says “curved upwards” or “curved downwards.”

You can use the second derivative to tell if a function is concave or convex.

11.1. Theorem. A function f is convex on some interval $a < x < b$ if and only if $f''(x) \geq 0$ for all x on that interval.

11.2. Theorem. A function f is concave on some interval $a < x < b$ if and only if the derivative $f'(x)$ is a nondecreasing function on that interval.

A proof using the Mean Value Theorem will be given in class.

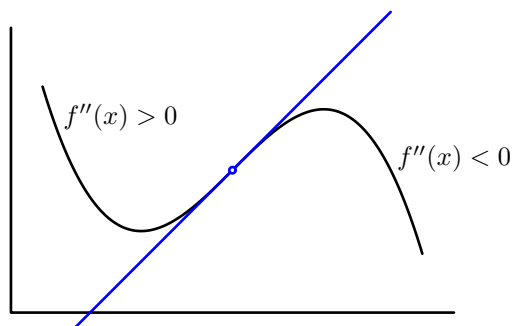


Figure 10. At an inflection point the tangent crosses the graph.

11.3. Example – the cubic function $f(x) = x^3 - x$. The second derivative of the function $f(x) = x^3 - x$ is

$$f''(x) = 6x$$

which is positive for $x > 0$ and negative for $x < 0$. Hence, in the graph in §6.3, the origin is an inflection point, and the piece of the graph where $x > 0$ is convex, while the piece where $x < 0$ is concave.

11.4. The second derivative test. In §7.3 we saw how you can tell if a stationary point is a local maximum or minimum by looking at the sign changes of $f'(x)$. There is another way of distinguishing between local maxima and minima which involves computing the second derivative.

11.5. Theorem. *If c is a stationary point for a function f , and if $f''(c) < 0$ then f has a local maximum at $x = c$.*

If $f''(c) > 0$ then f has a local minimum at c . The theorem doesn't say what happens when $f''(c) = 0$. In that case you must go back to checking the signs of the first derivative near the stationary point.

The basic reason why this theorem is true is that if c is a stationary point with $f''(c) > 0$ then " $f'(x)$ is increasing near $x = c$ " and hence $f'(x) < 0$ for $x < c$ and $f'(x) > 0$ for $x > c$. So the function f is decreasing for $x < c$ and increasing for $x > c$, and therefore it reaches a local minimum at $x = c$.

11.6. Example – that cubic function again. Consider the function $f(x) = x^3 - x$ from §6.3 and §11.3. We had found that this function has two stationary points, namely at $x = \pm\frac{1}{3}\sqrt{3}$. By looking at the sign of $f'(x) = 3x^2 - 1$ we concluded that $-\frac{1}{3}\sqrt{3}$ is a local maximum while $+\frac{1}{3}\sqrt{3}$ is a local minimum. Instead of looking at $f'(x)$ we could also have computed $f''(x)$ at $x = \pm\frac{1}{3}\sqrt{3}$ and applied the second derivative test. Here is how it goes:

Since $f''(x) = 6x$ we have

$$f''(-\frac{1}{3}\sqrt{3}) = -2\sqrt{3} < 0 \text{ and } f''(\frac{1}{3}\sqrt{3}) = 2\sqrt{3} > 0.$$

Therefore f has a local maximum at $-\frac{1}{3}\sqrt{3}$ and a local minimum at $\frac{1}{3}\sqrt{3}$.

11.7. When the second derivative test doesn't work. Usually the second derivative test will work, but sometimes a stationary point c has $f''(c) = 0$. In this case the second derivative test gives no information at all. The figure below shows you the graphs of three functions, all three of which have a stationary point at $x = 0$. In all three cases the second derivative vanishes at $x = 0$ so the second derivative test says nothing. As you can see, the stationary point can be a local maximum, a local minimum, or neither.

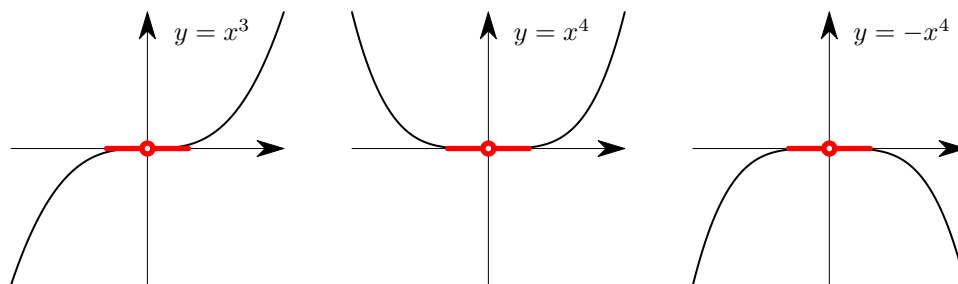


Figure 11. Three functions for which the second derivative test doesn't work.

12. Proofs of some of the theorems

12.1. Proof of the Mean Value Theorem. Let m be the slope of the chord connecting the points $(a, f(a))$ and $(b, f(b))$, i.e.

$$m = \frac{f(b) - f(a)}{b - a},$$

and consider the function

$$g(x) = f(x) - f(a) - m(x - a).$$

This function is continuous (since f is continuous), and g attains its maximum and minimum at two numbers c_{\min} and c_{\max} .

There are now two possibilities: either at least one of c_{\min} or c_{\max} is an interior point, or else both c_{\min} and c_{\max} are endpoints of the interval $a \leq x \leq b$.

Consider the first case: one of these two numbers is an interior point, i.e. if $a < c_{\min} < b$ or $a < c_{\max} < b$, then the derivative of g must vanish at c_{\min} or c_{\max} . If one has $g'(c_{\min}) = 0$, then one has

$$0 = g'(c_{\min}) = f'(c_{\min}) - m, \text{ i.e. } m = f'(c_{\min}).$$

The definition of m implies that one gets

$$f'(c_{\min}) = \frac{f(b) - f(a)}{b - a}.$$

If $g'(c_{\max}) = 0$ then one gets $m = f'(c_{\max})$ and hence

$$f'(c_{\max}) = \frac{f(b) - f(a)}{b - a}.$$

We are left with the remaining case, in which both c_{\min} and c_{\max} are end points. To deal with this case note that at the endpoints one has

$$g(a) = 0 \text{ and } g(b) = 0.$$

Thus the maximal and minimal values of g are both zero! This means that $g(x) = 0$ for all x , and thus that $g'(x) = 0$ for all x . Therefore we get $f'(x) = m$ for **all** x , and not just for some c .

12.2. Proof of Theorem 5.1. If f is a non-increasing function and if it is differentiable at some interior point a , then we must show that $f'(a) \geq 0$.

Since f is non-decreasing, one has $f(x) \geq f(a)$ for all $x > a$. Hence one also has

$$\frac{f(x) - f(a)}{x - a} \geq 0$$

for all $x > a$. Let $x \searrow a$, and you get

$$f'(a) = \lim_{x \searrow a} \frac{f(x) - f(a)}{x - a} \geq 0.$$

12.3. Proof of Theorem 5.2. Suppose f is a differentiable function on an interval $a < x < b$, and suppose that $f'(x) \geq 0$ on that interval. We must show that f is non-decreasing on that interval, i.e. we have to show that if $x_1 < x_2$ are two numbers in the interval (a, b) , then $f(x_1) \leq f(x_2)$. To prove this we use the Mean Value Theorem: given x_1 and x_2 the Mean Value Theorem hands us a number c with $x_1 < c < x_2$, and

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

We don't know where c is exactly, but it doesn't matter because we do know that wherever c is we have $f'(c) \geq 0$. Hence

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0.$$

Multiply with $x_2 - x_1$ (which we are allowed to do since $x_2 > x_1$ so $x_2 - x_1 > 0$) and you get

$$f(x_2) - f(x_1) \geq 0,$$

as claimed.

13. Exercises

218. What does the Intermediate Value Theorem say?

219. What does the Mean Value Theorem say?

220. Group Problem.

If $f(a) = 0$ and $f(b) = 0$ then there is a c between a and b such that $f'(c) = 0$. Show that this follows from the Mean Value Theorem. (Help! A proof! Relax: this one is not difficult. Make a drawing of the situation, then read the Mean Value Theorem again.)

221. What is a stationary point?

222. Group Problem.

How can you tell if a local maximum is a global maximum?

223. Group Problem.

If $f''(a) = 0$ then the graph of f has an inflection point at $x = a$. **True or False?**

224. What is an inflection point?

225. Give an example of a function for which $f'(0) = 0$ even though the function f has neither a local maximum or a local minimum at $x = 0$.

226. Group Problem.

Draw four graphs of functions, one for each of the following four combinations

$$\begin{array}{ll} f' > 0 \text{ and } f'' > 0 & f' > 0 \text{ and } f'' < 0 \\ f' < 0 \text{ and } f'' > 0 & f' < 0 \text{ and } f'' < 0 \end{array}$$

227. Group Problem.

Which of the following combinations are possible:

$$\begin{array}{l} f'(x) > 0 \text{ and } f''(x) = 0 \text{ for all } x \\ f'(x) = 0 \text{ and } f''(x) > 0 \text{ for all } x \end{array}$$

Sketch the graph of the following functions. You should

- (1) find where f , f' and f'' are positive or negative
- (2) find all stationary points
- (3) decide which stationary points are local maxima or minima
- (4) decide which local max/minima are in fact global max/minima
- (5) find all inflection points
- (6) find "horizontal asymptotes," i.e. compute the limits $\lim_{x \rightarrow \pm\infty} f(x)$ when appropriate.

228. $y = x^3 + 2x^2$

229. $y = x^3 - 4x^2$

230. $y = x^4 + 27x$

231. $y = x^4 - 27x$

232. $y = x^4 + 2x^2 - 3$

233. $y = x^4 - 5x^2 + 4$

234. $y = x^5 + 16x$

235. $y = x^5 - 16x$

236. $y = \frac{x}{x+1}$

237. $y = \frac{x}{1+x^2}$

238. $y = \frac{x^2}{1+x^2}$

239. $y = \frac{1+x^2}{1+x}$

240. $y = x + \frac{1}{x}$

241. $y = x - \frac{1}{x}$

242. $y = x^3 + 2x^2 + x$

243. $y = x^3 + 2x^2 - x$

244. $y = x^4 - x^3 - x$

245. $y = x^4 - 2x^3 + 2x$

246. $y = \sqrt{1+x^2}$

247. $y = \sqrt{1-x^2}$

248. $y = \sqrt[4]{1+x^2}$

249. $y = \frac{1}{1+x^4}$

The following functions are periodic, i.e. they satisfy $f(x+L) = f(x)$ for all x , where the constant L is called the period of the function. The graph of a periodic function repeats itself indefinitely to the left and to the right. It therefore has infinitely many (local) minima and maxima, and infinitely many inflections points. Sketch the graphs of the following functions as in the previous problem, but only list those "interesting points" that lie in the interval $0 \leq x < 2\pi$.

250. $y = \sin x$

251. $y = \sin x + \cos x$

252. $y = \sin x + \sin^2 x$

253. $y = 2 \sin x + \sin^2 x$

254. $y = 4 \sin x + \sin^2 x$

255. $y = 2 \cos x + \cos^2 x$

256. $y = \frac{4}{2 + \sin x}$

257. $y = (2 + \sin x)^2$

Find the domain and sketch the graphs of each of the following functions

258. $y = \arcsin x$

259. $y = \arctan x$

260. $y = 2 \arctan x - x$

261. $y = \arctan(x^2)$

262. $y = 3 \arcsin(x) - 5x$

263. $y = 6 \arcsin(x) - 10x^2$

264. In the following two problems it is not possible to solve the equation $f'(x) = 0$, but you can tell something from the second derivative.

(a) Show that the function $f(x) = x \arctan x$ is convex. Then sketch the graph of f .

(b) Show that the function $g(x) = x \arcsin x$ is convex. Then sketch the graph of g .

For each of the following functions use the derivative to decide if they are increasing, decreasing or neither on the indicated intervals

265. $f(x) = \frac{x}{1+x^2} \quad 10 < x < \infty$

266. $f(x) = \frac{2+x^2}{x^3-x} \quad 1 < x < \infty$

267. $f(x) = \frac{2+x^2}{x^3-x} \quad 0 < x < 1$

268. $f(x) = \frac{2+x^2}{x^3-x} \quad 0 < x < \infty$

14. Optimization Problems

Often a problem can be phrased as

For which value of x in the interval $a \leq x \leq b$ is $f(x)$ the largest?

In other words you are given a function f on an interval $[a, b]$ and you must find all global maxima of f on this interval.

If the function is continuous then according to theorem 8.1 there always is at least one x in the interval $[a, b]$ which maximizes $f(x)$.

If f is differentiable then we know what to do: any local maximum is either a stationary point or one of the end points a and b . Therefore you can find the global maxima by following this recipe:

- (1) Find all stationary points of f ;
- (2) Compute $f(x)$ at each stationary point you found in step (1);
- (3) Compute $f(a)$ and $f(b)$;
- (4) The global maxima are those stationary- or endpoints from steps (2) and (3) which have the largest function value.

Usually there is only one global maximum, but sometimes there can be more.

If you have to *minimize* rather than *maximize* a function, then you must look for global minima. The same recipe works (of course you should look for the smallest function value instead of the largest in step 4.)

The difficulty in optimization problems frequently lies not with the calculus part, but rather with setting up the problem. Choosing which quantity to call x and finding the function f is half the job.

14.1. Example – The rectangle with largest area and given perimeter. Which rectangle has the largest area, among all those rectangles for which the total length of the sides is 1?

Solution: If the sides of the rectangle have lengths x and y , then the total length of the sides is

$$L = x + x + y + y = 2(x + y)$$

and the area of the rectangle is

$$A = xy.$$

So are asked to find the largest possible value of $A = xy$ provided $2(x + y) = 1$. The lengths of the sides can also not be negative, so x and y must satisfy $x \geq 0$, $y \geq 0$.

We now want to turn this problem into a question of the form “maximize a function over some interval.” The quantity which we are asked to maximize is A , but it depends on two variables x and y instead of just one variable. However, the variables x and y are not independent since we are only allowed to consider rectangles with $L = 1$. From this equation we get

$$L = 1 \implies y = \frac{1}{2} - x.$$

Hence we must find the maximum of the quantity

$$A = xy = x\left(\frac{1}{2} - x\right)$$

The values of x which we are allowed to consider are only limited by the requirements $x \geq 0$ and $y \geq 0$, i.e. $x \leq \frac{1}{2}$. So we end up with this problem:

Find the maximum of the function $f(x) = x(\frac{1}{2} - x)$ on the interval $0 \leq x \leq \frac{1}{2}$.

Before we start computing anything we note that the function f is a polynomial so that it is differentiable, and hence continuous, and also that the interval $0 \leq x \leq \frac{1}{2}$ is closed. Therefore the theory guarantees that there is a maximum and our recipe will show us where it is.

The derivative is given by

$$f'(x) = \frac{1}{2} - 2x,$$

and hence the only stationary point is $x = \frac{1}{4}$. The function value at this point is

$$f\left(\frac{1}{4}\right) = \frac{1}{4}\left(\frac{1}{2} - \frac{1}{4}\right) = \frac{1}{16}.$$

At the endpoints one has $x = 0$ or $x = \frac{1}{2}$, which corresponds to a rectangle one of whose sides has length zero. The area of such rectangles is zero, and so this is not the maximal value we are looking for.

We conclude that the largest area is attained by the rectangle whose sides have lengths

$$x = \frac{1}{4}, \text{ and } y = \frac{1}{2} - \frac{1}{4} = \frac{1}{4},$$

i.e. by a square with sides $\frac{1}{4}$.

15. Exercises

269. By definition, the perimeter of a rectangle is the sum of the lengths of its four sides. Which rectangle, of all those whose perimeter is 1, has the smallest area? Which one has the largest area?

270. Which rectangle of area 100in^2 minimizes its height plus two times its length?

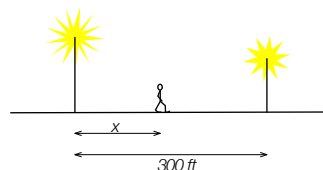
271. You have 1 yard of string from which you make a circular wedge with radius R and opening angle θ . Which choice of θ and R will give you the wedge with the largest area? Which choice leads to the smallest area?

[A circular wedge is the figure consisting of two radii of a circle and the arc connecting them. So the yard of string is used to form the two radii and the arc.]

272. Group Problem.

(The lamp post problem)

In a street two lamp posts are 300 feet apart. The light intensity at a distance d from the first lamp post is $1000/d^2$, the light intensity at distance d from the second (weaker) lamp post is $125/d^2$ (in both cases the light intensity is inversely proportional to the square of the distance to the light source).



The *combined light intensity* is the sum of the two light intensities coming from both lamp posts.

(a) If you are in between the lamp posts, at distance x feet from the stronger light, then give a formula for the combined light intensity coming from both lamp posts as a function of x .

(b) What is the darkest spot between the two lights, i.e. where is the combined light intensity the smallest?

273. (a) You have a sheet of metal with area 100 in^2 from which you are to make a cylindrical soup can. If r is the radius of the can and h its height, then which h and r will give you the can with the largest volume?

(b) If instead of making a plain cylinder you replaced the flat top and bottom of the cylinder with two spherical

caps, then (using the same 100in^2 of sheet metal), then which choice of radius and height of the cylinder give you the container with the largest volume?

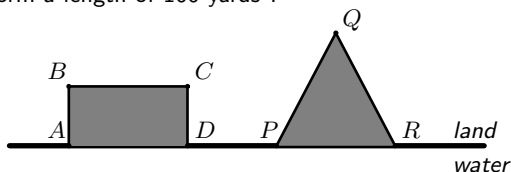
(c) Suppose you only replace the top of the cylinder with a spherical cap, and leave the bottom flat, then which choice of height and radius of the cylinder result in the largest volume?

274. A triangle has one vertex at the origin $O(0,0)$, another at the point $A(2a,0)$ and the third at $(a, a/(1+a^3))$. What are the largest and smallest areas this triangle can have if $0 \leq a < \infty$?

275. Group Problem.

Queen Dido's problem

According to tradition Dido was the founder and first Queen of Carthage. When she arrived on the north coast of Africa (~800BC) the locals allowed her to take as much land as could be enclosed with the hide of one ox. She cut the hide into thin strips and put these together to form a length of 100 yards¹.



(a) If Dido wanted a rectangular region, then how wide should she choose it to enclose as much area as possible (the coastal edge of the boundary doesn't count, so in this problem the length $AB + BC + CD$ is 100 yards.)

(b) If Dido chose a region in the shape of an isosceles triangle PQR , then how wide should she make it to maximize its area (again, don't include the coast in the perimeter: $PQ + QR$ is 100 yards long, and $PQ = QR$.)

276. The product of two numbers x, y is 16. We know $x \geq 1$ and $y \geq 1$. What is the greatest possible sum of the two numbers?

277. What are the smallest and largest values that $(\sin x)(\sin y)$ can have if $x + y = \pi$ and if x and y are both nonnegative?

278. What are the smallest and largest values that $(\cos x)(\cos y)$ can have if $x + y = \frac{\pi}{2}$ and if x and y are both nonnegative?

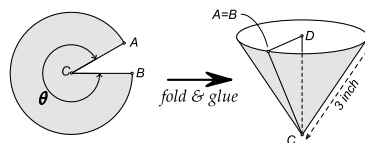
279. (a) What are the smallest and largest values that $\tan x + \tan y$ can have if $x + y = \frac{\pi}{2}$ and if x and y are both nonnegative?

(b) What are the smallest and largest values that $\tan x + 2 \tan y$ can have if $x + y = \frac{\pi}{2}$ and if x and y are both nonnegative?

280. The cost per hour of fuel to run a locomotive is $v^2/25$ dollars, where v is speed (in miles per hour), and other costs are \$100 per hour regardless of speed. What is the speed that minimizes cost per mile?

281. Group Problem.

Josh is in need of coffee. He has a circular filter with 3 inch radius. He cuts out a wedge and glues the two edges AC and BC together to make a conical filter to hold the ground coffee. The volume V of the coffee cone depends the angle θ of the piece of filter paper Josh made.



(a) Find the volume in terms of the angle θ . (Hint: how long is the circular arc AB on the left? How long is the circular top of the cone on the right? If you know that you can find the radius $AD = BD$ of the top of the cone, and also the height CD of the cone.)

(b) Which angle θ maximizes the volume V ?

¹I made that number up. For the rest start at <http://en.wikipedia.org/wiki/Dido>

Exponentials and Logarithms (naturally)

In this chapter we first recall some facts about exponentials (x^y with $x > 0$ and y arbitrary): they should be familiar from algebra, or “precalculus.” What is new is perhaps the definition of x^y when y is not a fraction: e.g., $2^{3/4}$ is the 4th root of the third power of 2 ($\sqrt[4]{2^3}$), but what is $2^{\sqrt{2}}$?

Then we ask “what is the derivative of $f(x) = a^x$?” The answer leads us to the famous number $e \approx 2.718\,281\,828\,459\,045\,235\,360\,287\,471\,352\,662\,497\,757\,247\,093\,699\,95\dots$.

Finally, we compute the derivative of $f(x) = \log_a x$, and we look at things that “grow exponentially.”

1. Exponents

Here we go over the definition of x^y when x and y are arbitrary real numbers, with $x > 0$.

For any real number x and any positive integer $n = 1, 2, 3, \dots$ one defines

$$x^n = \overbrace{x \cdot x \cdot \dots \cdot x}^{n \text{ times}}$$

and, if $x \neq 0$,

$$x^{-n} = \frac{1}{x^n}.$$

One defines $x^0 = 1$ for any $x \neq 0$.

To define $x^{p/q}$ for a general fraction $\frac{p}{q}$ one must assume that the number x is positive. One then defines

$$(37) \quad x^{p/q} = \sqrt[q]{x^p}.$$

This does not tell us how to define x^a if the exponent a is not a fraction. One can define x^a for irrational numbers a by taking limits. For example, to define $2^{\sqrt{2}}$, we look at the sequence of numbers you get by truncating the decimal expansion of $\sqrt{2}$, i.e.

$$a_1 = 1, \quad a_2 = 1.4 = \frac{14}{10}, \quad a_3 = 1.41 = \frac{141}{100}, \quad a_4 = 1.414 = \frac{1414}{1000}, \quad \dots$$

Each a_n is a fraction, so that we know what 2^{a_n} is, e.g. $2^{a_4} = \sqrt[1000]{2^{1414}}$. Our definition of $2^{\sqrt{2}}$ then is

$$2^{\sqrt{2}} = \lim_{n \rightarrow \infty} 2^{a_n},$$

i.e. we define $2^{\sqrt{2}}$ as the limit of the sequence of numbers

$$2, \sqrt[10]{2^{14}}, \sqrt[100]{2^{141}}, \sqrt[1000]{2^{1414}}, \dots$$

(See table 1.)

Here one ought to prove that this limit exists, and that its value does not depend on the particular choice of numbers a_n tending to a . We will not go into these details in this course.

It is shown in precalculus texts that the exponential functions satisfy the following properties:

$$(38) \quad \boxed{x^a x^b = x^{a+b}, \quad \frac{x^a}{x^b} = x^{a-b}, \quad (x^a)^b = x^{ab}}$$

provided a and b are fractions. One can show that these properties still hold if a and b are real numbers (not necessarily fractions.) Again, we won’t go through the proofs here.

| x | 2^x |
|--------------|----------------|
| 1.0000000000 | 2.000000000000 |
| 1.4000000000 | 2.639015821546 |
| 1.4100000000 | 2.657371628193 |
| 1.4140000000 | 2.664749650184 |
| 1.4142000000 | 2.665119088532 |
| 1.4142100000 | 2.665137561794 |
| 1.4142130000 | 2.665143103798 |
| 1.4142135000 | 2.665144027466 |
| \vdots | \vdots |

Table 1. Approximating $2^{\sqrt{2}}$. Note that as x gets closer to $\sqrt{2}$ the quantity 2^x appears to converge to some number. This limit is our definition of $2^{\sqrt{2}}$.

Now instead of considering x^a as a function of x we can pick a positive number a and consider the function $f(x) = a^x$. This function is defined for all real numbers x (as long as the base a is positive.).

1.1. The trouble with powers of negative numbers. The cube root of a negative number is well defined. For instance $\sqrt[3]{-8} = -2$ because $(-2)^3 = -8$. In view of the definition (37) of $x^{p/q}$ we can write this as

$$(-8)^{1/3} = \sqrt[3]{(-8)^1} = \sqrt[3]{-8} = -2.$$

But there is a problem: since $\frac{2}{6} = \frac{1}{3}$ you would think that $(-8)^{2/6} = (-8)^{1/3}$. However our definition (37) tells us that

$$(-8)^{2/6} = \sqrt[6]{(-8)^2} = \sqrt[6]{+64} = +2.$$

Another example:

$$(-4)^{1/2} = \sqrt{-4} \text{ is not defined}$$

but, even though $\frac{1}{2} = \frac{2}{4}$,

$$(-4)^{2/4} = \sqrt[4]{(-4)^2} = \sqrt[4]{+16} = 2 \text{ is defined.}$$

There are two ways out of this mess:

- (1) avoid taking fractional powers of negative numbers
- (2) when you compute $x^{p/q}$ first simplify the fraction by removing common divisors of p and q .

The safest is just not to take fractional powers of negative numbers.

Given that fractional powers of negative numbers cause all these headaches it is not surprising that we didn't try to define x^a for negative x if a is irrational. For example, $(-8)^\pi$ is not defined¹.

2. Logarithms

Briefly, $y = \log_a x$ is the inverse function to $y = a^x$. This means that, by definition,

$$y = \log_a x \iff x = a^y.$$

In other words, $\log_a x$ is the answer to the question "for which number y does one have $x = a^y$?" The number $\log_a x$ is called **the logarithm with base a of x** . In this definition both a and x must be positive.

For instance,

$$2^3 = 8, \quad 2^{1/2} = \sqrt{2}, \quad 2^{-1} = \frac{1}{2}$$

so

$$\log_2 8 = 3, \quad \log_2(\sqrt{2}) = \frac{1}{2}, \quad \log_2 \frac{1}{2} = -1.$$

¹There is a definition of $(-8)^\pi$ which uses complex numbers. You will see this next semester if you take math 222.

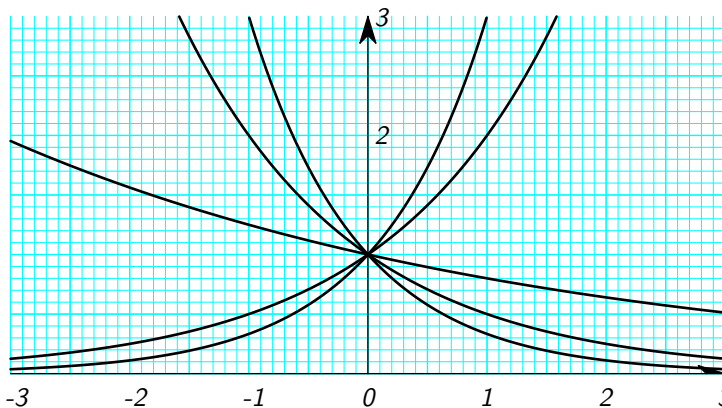


Figure 1. The graphs of $y = 2^x, 3^x, (1/2)^x, (0.1)^x$ and $y = (4/5)^x$. The graphs are purposely not labeled: can you figure out which is which?

Also:

$\log_2(-3)$ doesn't exist

because there is no number y for which $2^y = -3$ (2^y is always positive) and

$\log_{-3} 2$ doesn't exist either

because $y = \log_{-3} 2$ would have to be some real number which satisfies $(-3)^y = 2$, and we don't take non-integer powers of negative numbers.

3. Properties of logarithms

In general one has

$$\log_a a^x = x, \text{ and } a^{\log_a x} = x.$$

There is a subtle difference between these formulas: the first one holds for all real numbers x , but the second only holds for $x > 0$, since $\log_a x$ doesn't make sense for $x \leq 0$.

Again, one finds the following formulas in precalculus texts:

$$(39) \quad \begin{array}{l} \log_a xy = \log_a x + \log_a y \\ \log_a \frac{x}{y} = \log_a x - \log_a y \\ \log_a x^y = y \log_a x \\ \log_a x = \frac{\log_b x}{\log_b a} \end{array}$$

They follow from (38).

4. Graphs of exponential functions and logarithms

Figure 1 shows the graphs of some exponential functions $y = a^x$ with different values of a , and figure 2 shows the graphs of $y = \log_2 x$, $y = \log_3 x$, $\log_{1/2} x$, $\log_{1/3} x$ and $y = \log_{10} x$. Can you tell which is which? (Yes, you can.)

From algebra/precalc recall:

If $a > 1$ then $f(x) = a^x$ is an increasing function.

and

If $0 < a < 1$ then $f(x) = a^x$ is a decreasing function.

In other words, for $a > 1$ it follows from $x_1 < x_2$ that $a^{x_1} < a^{x_2}$; if $0 < a < 1$, then $x_1 < x_2$ implies $a^{x_1} > a^{x_2}$.

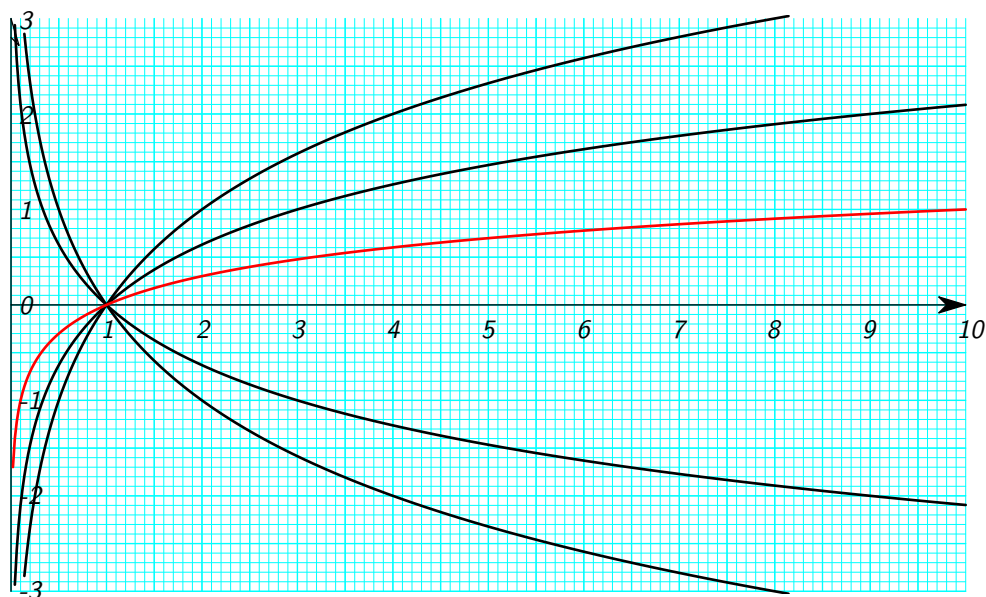


Figure 2. Graphs of some logarithms. Each curve is the graph of a function $y = \log_a x$ for various values of $a > 0$. Can you tell what a is for each graph?

5. The derivative of a^x and the definition of e

To begin, we try to differentiate the function $y = 2^x$:

$$\begin{aligned} \frac{d2^x}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{2^{x+\Delta x} - 2^x}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{2^x 2^{\Delta x} - 2^x}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} 2^x \frac{2^{\Delta x} - 1}{\Delta x} \\ &= 2^x \lim_{\Delta x \rightarrow 0} \frac{2^{\Delta x} - 1}{\Delta x}. \end{aligned}$$

So if we assume that the limit

$$\lim_{\Delta x \rightarrow 0} \frac{2^{\Delta x} - 1}{\Delta x} = C$$

exists then we have

$$(40) \quad \frac{d2^x}{dx} = C2^x.$$

On your calculator you can compute $\frac{2^{\Delta x} - 1}{\Delta x}$ for smaller and smaller values of Δx , which leads you to suspect that the limit actually exists, and that $C \approx 0.693147 \dots$. One can in fact prove that the limit exists, but we will not do this here.

Once we know (40) we can compute the derivative of a^x for any other positive number a . To do this we write $a = 2^{\log_2 a}$, and hence

$$a^x = (2^{\log_2 a})^x = 2^{x \cdot \log_2 a}.$$

By the chain rule we therefore get

$$\begin{aligned}\frac{da^x}{dx} &= \frac{d2^{x \cdot \log_2 a}}{dx} \\ &= C 2^{x \cdot \log_2 a} \frac{dx \cdot \log_2 a}{dx} \\ &= (C \log_2 a) 2^{x \cdot \log_2 a} \\ &= (C \log_2 a) a^x.\end{aligned}$$

So the derivative of a^x is just some constant times a^x , the constant being $C \log_2 a$. This is essentially our formula for the derivative of a^x , but one can make the formula look nicer by introducing a special number, namely, we define

$$e = 2^{1/C} \text{ where } C = \lim_{\Delta x \rightarrow 0} \frac{2^{\Delta x} - 1}{\Delta x}.$$

One has

$$e \approx 2.718\,281\,818\,459 \dots$$

This number is special because if you set $a = e$, then

$$C \log_2 a = C \log_2 e = C \log_2 2^{1/C} = C \cdot \frac{1}{C} = 1,$$

and therefore the derivative of the function $y = e^x$ is

$$(41) \quad \boxed{\frac{de^x}{dx} = e^x.}$$

Read that again: the function e^x is its own derivative!

The logarithm with base e is called the **Natural Logarithm**, and is written

$$\ln x = \log_e x.$$

Thus we have

$$(42) \quad \boxed{e^{\ln x} = x \quad \ln e^x = x}$$

where the second formula holds for all real numbers x but the first one only makes sense for $x > 0$.

For any positive number a we have $a = e^{\ln a}$, and also

$$a^x = e^{x \ln a}.$$

By the chain rule you then get

$$(43) \quad \boxed{\frac{da^x}{dx} = a^x \ln a.}$$

6. Derivatives of Logarithms

Since the natural logarithm is the inverse function of $f(x) = e^x$ we can find its derivative by implicit differentiation. Here is the computation (which you should do yourself)

The function $f(x) = \log_a x$ satisfies

$$a^{f(x)} = x$$

Differentiate both sides, and use the chain rule on the left,

$$(\ln a) a^{f(x)} f'(x) = 1.$$

Then solve for $f'(x)$ to get

$$f'(x) = \frac{1}{(\ln a) a^{f(x)}}.$$

Finally we remember that $a^{f(x)} = x$ which gives us the derivative of a^x

$$\frac{da^x}{dx} = \frac{1}{x \ln a}.$$

In particular, the natural logarithm has a very simple derivative, namely, since $\ln e = 1$ we have

$$(44) \quad \boxed{\frac{d \ln x}{dx} = \frac{1}{x}}.$$

7. Limits involving exponentials and logarithms

7.1. Theorem. *Let r be any real number. Then, if $a > 1$,*

$$\lim_{x \rightarrow \infty} x^r a^{-x} = 0,$$

i.e.

$$\lim_{x \rightarrow \infty} \frac{x^r}{a^x} = 0.$$

This theorem says that *any exponential will beat any power of x as $x \rightarrow \infty$* . For instance, as $x \rightarrow \infty$ both x^{1000} and $(1.001)^x$ go to infinity, but

$$\lim_{x \rightarrow \infty} \frac{x^{1000}}{(1.001)^x} = 0,$$

so, in the long run, for very large x , 1.001^x will be much larger than 1000^x .

PROOF WHEN $a = e$. We want to show $\lim_{x \rightarrow \infty} x^r e^{-x} = 0$. To do this consider the function $f(x) = x^{r+1} e^{-x}$. Its derivative is

$$f'(x) = \frac{dx^{r+1}e^{-x}}{dx} = ((r+1)x^r - x^{r+1})e^{-x} = (r+1-x)x^r e^{-x}.$$

Therefore $f'(x) < 0$ for $x > r+1$, i.e. $f(x)$ is decreasing for $x > r+1$. It follows that $f(x) < f(r+1)$ for all $x > r+1$, i.e.

$$x^{r+1} e^{-x} < (r+1)^{r+1} e^{-(r+1)} \text{ for } x > r+1.$$

Divide by x , abbreviate $A = (r+1)^{r+1} e^{-(r+1)}$, and we get

$$0 < x^r e^{-x} < \frac{A}{x} \text{ for all } x > r+1.$$

The Sandwich Theorem implies that $\lim_{x \rightarrow \infty} x^r e^{-x} = 0$, which is what we had promised to show. □

Here are some related limits:

$$\begin{aligned} a > 1 &\implies \lim_{x \rightarrow \infty} \frac{a^x}{x^r} = \infty \quad (D.N.E.) \\ m > 0 &\implies \lim_{x \rightarrow \infty} \frac{\ln x}{x^m} = 0 \\ m > 0 &\implies \lim_{x \rightarrow 0} x^m \ln x = 0 \end{aligned}$$

The second limit says that even though $\ln x$ becomes infinitely large as $x \rightarrow \infty$, it is always much less than any power x^m with $m > 0$ real. To prove it you set $x = e^t$ and then $t = s/m$, which leads to

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x^m} \stackrel{x=e^t}{=} \lim_{t \rightarrow \infty} \frac{t}{e^{mt}} \stackrel{t=s/m}{=} \frac{1}{m} \lim_{t \rightarrow \infty} \frac{s}{e^s} = 0.$$

The third limit follows from the second by substituting $x = 1/y$ and using $\ln \frac{1}{x} = -\ln x$.

8. Exponential growth and decay

A quantity X which depends on time t is said to grow or decay exponentially if it is given by

$$(45) \quad X(t) = X_0 e^{kt}.$$

The constant X_0 is the value of $X(t)$ at time $t = 0$ (sometimes called “the initial value of X ”).

The derivative of an exponentially growing quantity, i.e. its rate of change with time, is given by $X'(t) = X_0 k e^{kt}$ so that

$$(46) \quad \frac{dX(t)}{dt} = kX(t).$$

In words, *for an exponentially growing quantity the rate of change is always proportional to the quantity itself*. The proportionality constant is k and is sometimes called “the relative growth rate.”

This property of exponential functions completely describes them, by which I mean that any function which satisfies (46) automatically satisfies (45). To see that this is true, suppose you have a function $X(t)$ for which $X'(t) = kX(t)$ holds at all times t . Then

$$\begin{aligned} \frac{dX(t)e^{-kt}}{dt} &= X(t) \frac{de^{-kt}}{dt} + \frac{dX(t)}{dt} e^{-kt} \\ &= -kX(t)e^{-kt} + X'(t)e^{-kt} \\ &= (X'(t) - kX(t))e^{-kt} \\ &= 0. \end{aligned}$$

It follows that $X(t)e^{-kt}$ does not depend on t . At $t = 0$ one has

$$X(t)e^{-kt} = X(0)e^0 = X_0$$

and therefore we have

$$X(t)e^{-kt} = X_0 \text{ for all } t.$$

Multiply with e^{kt} and we end up with

$$X(t) = X_0 e^{kt}.$$

8.1. Half time and doubling time. If $X(t) = X_0 e^{kt}$ then one has

$$X(t+T) = X_0 e^{kt+kT} = X_0 e^{kt} e^{kT} = e^{kT} X(t).$$

In words, after time T goes by an exponentially growing (decaying) quantity changes by a factor e^{kT} . If $k > 0$, so that the quantity is actually growing, then one calls

$$T = \frac{\ln 2}{k}$$

the **doubling time** for X because $X(t)$ changes by a factor $e^{kT} = e^{\ln 2} = 2$ every T time units: $X(t)$ doubles every T time units.

If $k < 0$ then $X(t)$ is decaying and one calls

$$T = \frac{\ln 2}{-k}$$

the **half life** because $X(t)$ is reduced by a factor $e^{kT} = e^{-\ln 2} = \frac{1}{2}$ every T time units.

8.2. Determining X_0 and k . The general exponential growth/decay function (45) contains only two constants, X_0 and k , and if you know the values of $X(t)$ at two different times then you can compute these constants.

Suppose that you know

$$X_1 = X(t_1) \text{ and } X_2 = X(t_2).$$

Then we have

$$X_0 e^{kt_1} = X_1 \text{ and } X_2 = X_0 e^{kt_2}$$

in which t_1, t_2, X_1, X_2 are given and k and X_0 are unknown. One first finds k from

$$\frac{X_1}{X_2} = \frac{X_0 e^{kt_1}}{X_0 e^{kt_2}} = e^{k(t_1 - t_2)} \implies \ln \frac{X_1}{X_2} = k(t_1 - t_2)$$

which implies

$$k = \frac{\ln X_1 - \ln X_2}{t_1 - t_2}.$$

Once you have computed k you can find X_0 from

$$X_0 = \frac{X_1}{e^{kt_1}} = \frac{X_2}{e^{kt_2}}.$$

(both expressions should give the same result.)

9. Exercises

Sketch the graphs of the following functions.

(Hint for some of these: if you have to solve something like $e^{4x} - 3e^{3x} + e^x = 0$, then call $w = e^x$, and you get a polynomial equation for w , namely $w^4 - 3w^3 + w = 0$.)

282. $y = e^x$

283. $y = e^{-x}$

284. $y = e^x + e^{-2x}$

285. $y = e^{3x} - 4e^x$

286. $y = \frac{e^x}{1 + e^x}$

287. $y = \frac{2e^x}{1 + e^{2x}}$

288. $y = xe^{-x}$

289. $y = \sqrt{x}e^{-x/4}$

290. $y = x^2 e^{x+2}$

291. $y = e^{x/2} - x$

292. $y = \ln \sqrt{x}$

293. $y = \ln \frac{1}{x}$

294. $y = x \ln x$

295. $y = \frac{-1}{\ln x} \quad (0 < x < \infty, x \neq 1)$

296. $y = (\ln x)^2 \quad (x > 0)$

297. $y = \frac{\ln x}{x} \quad (x > 0)$

298. $y = \ln \sqrt{\frac{1+x}{1-x}} \quad (|x| < 1)$

299. $y = \ln(1 + x^2)$

300. $y = \ln(x^2 - 3x + 2) \quad (x > 2)$

301. $y = \ln \cos x \quad (|x| < \frac{\pi}{2})$

302. The function $f(x) = e^{-x^2}$ plays a central role in statistics and its graph is called **the bell curve** (because of its shape). Sketch the graph of f .

303. Sketch the part of the graph of the function

$$f(x) = e^{-\frac{1}{x}}$$

with $x > 0$.

Find the limits

$$\lim_{x \searrow 0} \frac{f(x)}{x^n} \text{ and } \lim_{x \rightarrow \infty} f(x)$$

where n can be any positive integer (hint: substitute $x = \dots$?)

304. A **damped oscillation** is a function of the form

$$f(x) = e^{-ax} \cos bx \text{ or } f(x) = e^{-ax} \sin bx$$

where a and b are constants.

Sketch the graph of $f(x) = e^{-x} \sin 10x$ (i.e. find zeroes, local max and mins, inflection points) and draw (with pencil on paper) the piece of the graph with $0 \leq x \leq 2\pi$.

This function has many local maxima and minima. What is the ratio between the function values at two consecutive local maxima? (Hint: the answer does not depend on which pair of consecutive local maxima you consider.)

305. Find the inflection points on the graph of $f(x) = (1+x)\ln x$ ($x > 0$).

306. (a) If x is large, which is bigger: 2^x or x^2 ?
 (b) The graphs of $f(x) = x^2$ and $g(x) = 2^x$ intersect at $x = 2$ (since $2^2 = 2^2$). How many more intersections do these graphs have (with $-\infty < x < \infty$)?

Find the following limits.

307. $\lim_{x \rightarrow \infty} \frac{e^x - 1}{e^x + 1}$

308. $\lim_{x \rightarrow \infty} \frac{e^x - x^2}{e^x + x}$

309. $\lim_{x \rightarrow \infty} \frac{2^x}{3^x - 2^x}$

310. $\lim_{x \rightarrow \infty} \frac{e^x - x^2}{e^{2x} + e^{-x}}$

311. $\lim_{x \rightarrow \infty} \frac{e^{-x} - e^{-x/2}}{\sqrt{e^x + 1}}$

312. $\lim_{x \rightarrow \infty} \frac{\sqrt{x + e^{4x}}}{e^{2x} + x}$

313. $\lim_{x \rightarrow \infty} \frac{e^{\sqrt{x}}}{\sqrt{e^x + 1}}$

314. $\lim_{x \rightarrow \infty} \ln(1+x) - \ln x$

315. $\lim_{x \rightarrow \infty} \frac{\ln x}{\ln x^2}$

316. $\lim_{x \rightarrow 0} x \ln x$

317. $\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt{x} + \ln x}$

318. $\lim_{x \rightarrow 0} \frac{\ln x}{\sqrt{x} + \ln x}$

319. Find the tenth derivative of xe^x .

320. For which real number x is $2^x - 3^x$ the largest?

321. Find $\frac{dx^x}{dx}$, $\frac{dx^{x^x}}{dx}$, and $\frac{d(x^x)^x}{dx}$.

Hint: $x^x = e^{x \ln x}$.

322. Group Problem.

About logarithmic differentiation:

(a) Let $y = (x+1)^2(x+3)^4(x+5)^6$ and $u = \ln y$. Find du/dx . Hint: Use the fact that \ln converts multiplication to addition before you differentiate. It will simplify the calculation.

(b) Check that the derivative of $\ln u(x)$ is the logarithmic derivative of the function u (as defined in the exercises following §25, chapter 4.)

323. After 3 days a sample of radon-222 decayed to 58% of its original amount.

(a) What is the half life of radon-222?

(b) How long would it take the sample to decay to 10% of its original amount?

324. Polonium-210 has a half life of 140 days.

(a) If a sample has a mass of 200 mg find a formula for the mass that remains after t days.

(b) Find the mass after 100 days.

(c) When will the mass be reduced to 10 mg?

(d) Sketch the graph of the mass as a function of time.

325. Current agricultural experts believe that the world's farms can feed about 10 billion people. The 1950 world population was 2.517 billion and the 1992 world population was 5.4 billion. When can we expect to run out of food?

326. Group Problem.

The ACME company runs two ads on Sunday mornings. One says that "when this baby is old enough to vote, the world will have one billion new mouths to feed" and the other says "in thirty six years, the world will have to set eight billion places at the table." What does ACME think the population of the world is at present? How fast does ACME think the population is increasing? Use units of billions of people so you can write 8 instead of 8,000,000,000. (Hint: $36 = 2 \times 18$.)

327. The population of California grows exponentially at an instantaneous rate of 2% per year. The population of California on January 1, 2000 was 20,000,000.

(a) Write a formula for the population $N(t)$ of California t years after January 1, 2000.

(b) Each Californian consumes pizzas at the rate of 70 pizzas per year. At what rate is California consuming pizzas t years after 1990?

(c) How many pizzas were consumed in California from January 1, 2005 to January 1, 2009?

328. The population of the country of Farfaraway grows exponentially.

(a) If its population in the year 1980 was 1,980,000 and its population in the year 1990 was 1,990,000, what is its population in the year 2000?

(b) How long will it take the population to double? (Your answer may be expressed in terms of exponentials and natural logarithms.)

329. The hyperbolic functions are defined by

$$\sinh x = \frac{e^x - e^{-x}}{2},$$

$$\cosh x = \frac{e^x + e^{-x}}{2},$$

$$\tanh x = \frac{\sinh x}{\cosh x}.$$

(a) Prove the following identities

$$\cosh^2 x - \sinh^2 x = 1$$

$$\cosh 2x = \cosh^2 x + \sinh^2 x$$

$$\sinh 2x = 2 \sinh x \cosh x.$$

(b) Show that

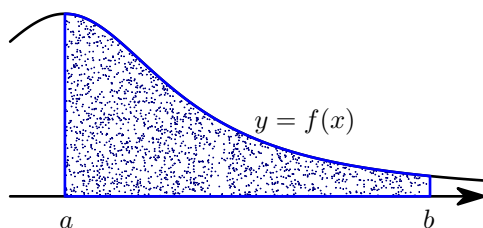
$$\begin{aligned}\frac{d \sinh x}{dx} &= \cosh x, \\ \frac{d \cosh x}{dx} &= \sinh x, \\ \frac{d \tanh x}{dx} &= \frac{1}{\cosh^2 x}.\end{aligned}$$

(c) Sketch the graphs of the three hyperbolic functions.

The Integral

In this chapter we define the integral of a function on some interval $[a, b]$. The most common interpretation of the integral is in terms of the area under the graph of the given function, so that is where we begin.

1. Area under a Graph



Let f be a function which is defined on some interval $a \leq x \leq b$ and assume it is positive, i.e. assume that its graph lies above the x axis. *How large is the area of the region caught between the x axis, the graph of $y = f(x)$ and the vertical lines $y = a$ and $y = b$?*

You can try to compute this area by approximating the region with many thin rectangles. Look at figure 1 before you read on. To make the approximating region you choose a *partition* of the interval $[a, b]$, i.e. you pick numbers $x_1 < \dots < x_n$ with

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b.$$

These numbers split the interval $[a, b]$ into n sub-intervals

$$[x_0, x_1], \quad [x_1, x_2], \quad \dots, \quad [x_{n-1}, x_n]$$

whose lengths are

$$\Delta x_1 = x_1 - x_0, \quad \Delta x_2 = x_2 - x_1, \quad \dots, \quad \Delta x_n = x_n - x_{n-1}.$$

In each interval we choose a point c_k , i.e. in the first interval we choose $x_0 \leq c_1 \leq x_1$, in the second interval we choose $x_1 \leq c_2 \leq x_2$, \dots , and in the last interval we choose some number $x_{n-1} \leq c_n \leq x_n$. See figure 1.

We then define n rectangles: the base of the k^{th} rectangle is the interval $[x_{k-1}, x_k]$ on the x -axis, while its height is $f(c_k)$ (here k can be any integer from 1 to n .)

The area of the k^{th} rectangle is of course the product of its height and width, i.e. its area is $f(c_k)\Delta x_k$. Adding these we see that the total area of the rectangles is

$$(47) \quad R = f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \dots + f(c_n)\Delta x_n.$$

This kind of sum is called a **Riemann sum**.

If the partition is sufficiently fine then one would expect this sum, i.e. the total area of all rectangles to be a good approximation of the area of the region under the graph. Replacing the partition by a finer partition, with more division points, should improve the approximation. So you would expect the area to be the limit of Riemann-sums like R “as the partition becomes finer and finer.” A precise formulation of the definition goes like this:

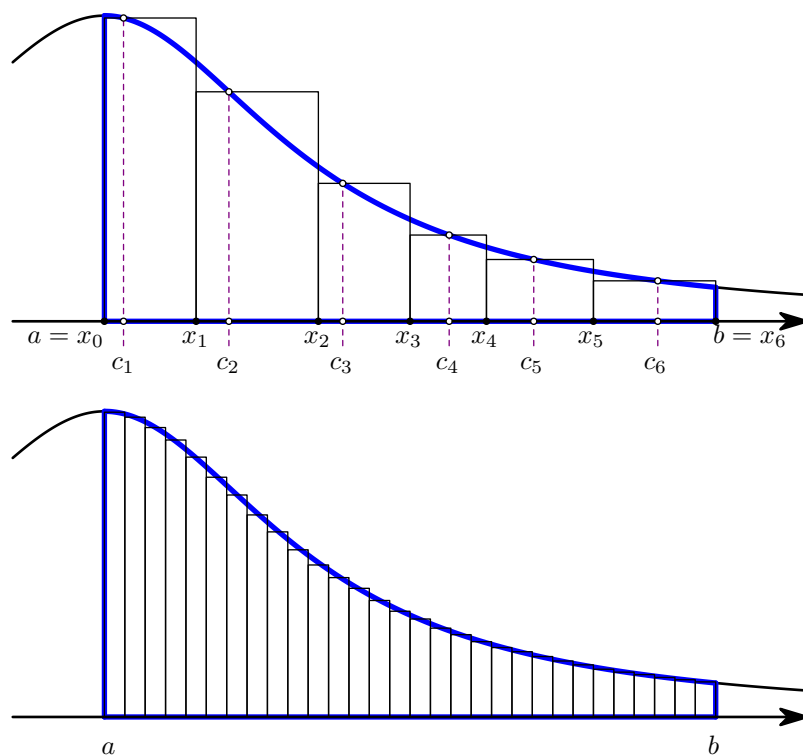


Figure 1. TOP. A Riemann sum in which the interval $a < x < b$ has been cut up into six smaller intervals. In each of those intervals a point c_i has been chosen at random, and the resulting rectangles with heights $f(c_1), \dots, f(c_6)$ were drawn. The total area under the graph of the function is roughly equal to the total area of the rectangles.

BOTTOM. Refining the partition. After adding more partition points the combined area of the rectangles will be a better approximation of the area under the graph of the function f .

1.1. Definition. If f is a function defined on an interval $[a, b]$, then we say that

$$\int_a^b f(x) dx = I,$$

i.e. the integral of “ $f(x)$ from $x = a$ to b ” equals I , if for every $\varepsilon > 0$ one can find a $\delta > 0$ such that

$$\left| f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \dots + f(c_n)\Delta x_n - I \right| < \varepsilon$$

holds for every partition all of whose intervals have length $\Delta x_k < \delta$.

2. When f changes its sign

If the function f is not necessarily positive everywhere in the interval $a \leq x \leq b$, then we still define the integral in exactly the same way: as a limit of Riemann sums whose mesh size becomes smaller and smaller. However the interpretation of the integral as “the area of the region between the graph and the x -axis” has a twist to it.

Let f be some function on an interval $a \leq x \leq b$, and form the Riemann sum

$$R = f(c_1)\Delta x_1 + f(c_2)\Delta x_2 + \dots + f(c_n)\Delta x_n$$

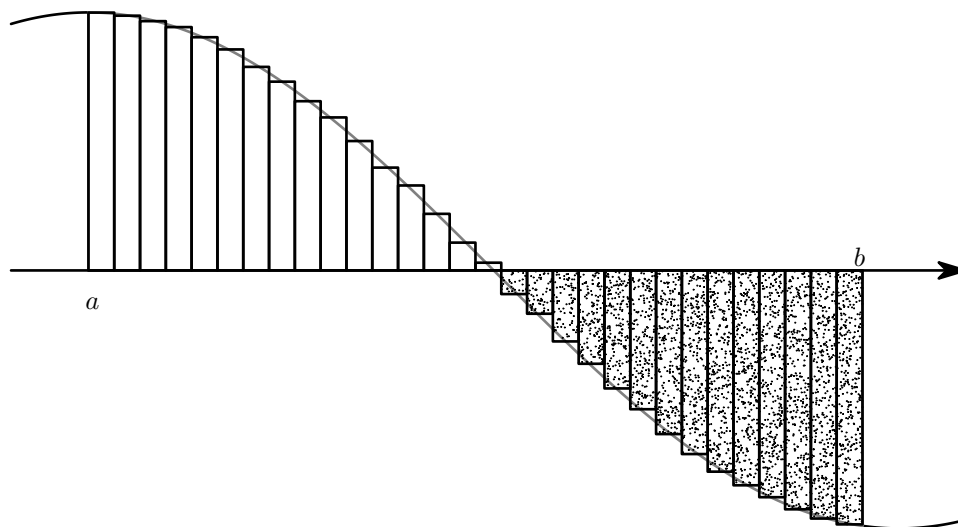


Figure 2. Illustrating a Riemann sum for a function whose sign changes. Always remember this: **AREAS ARE ALWAYS POSITIVE NUMBERS.** The Riemann-sum corresponding to this picture is the total area of the rectangles above the x -axis *minus* the total area of the rectangles below the x -axis.

that goes with some partition, and some choice of c_k .

When f can be positive or negative, then the terms in the Riemann sum can also be positive or negative. If $f(c_k) > 0$ then the quantity $f(c_k)\Delta x_k$ is the area of the corresponding rectangle, but if $f(c_k) < 0$ then $f(c_k)\Delta x_k$ is a negative number, namely *minus* the area of the corresponding rectangle. The Riemann sum is therefore the area of the rectangles above the x -axis *minus* the area below the axis and above the graph.

Taking the limit over finer and finer partitions, we conclude that

$$\int_a^b f(x)dx = \begin{array}{l} \text{area above the } x\text{-axis, below the graph} \\ \textit{minus} \text{ the area below the } x\text{-axis, above the graph.} \end{array}$$

3. The Fundamental Theorem of Calculus

3.1. Definition. A function F is called an **antiderivative** of f on the interval $[a, b]$ if one has $F'(x) = f(x)$ for all x with $a < x < b$.

For instance, $F(x) = \frac{1}{2}x^2$ is an antiderivative of $f(x) = x$, but so is $G(x) = \frac{1}{2}x^2 + 2008$.

3.2. Theorem. If f is a function whose integral $\int_a^b f(x)dx$ exists, and if F is an antiderivative of f on the interval $[a, b]$, then one has

$$(48) \quad \int_a^b f(x)dx = F(b) - F(a).$$

(a proof was given in lecture.)

Because of this theorem the expression on the right appears so often that various abbreviations have been invented. We will abbreviate

$$F(b) - F(a) \stackrel{\text{def}}{=} [F(x)]_{x=a}^b = [F(x)]_a^b.$$

3.3. Terminology.

In the integral

$$\int_a^b f(x) dx$$

the numbers a and b are called the *bounds of the integral*, the function $f(x)$ which is being integrated is called *the integrand*, and the variable x is *integration variable*.

The integration variable is a *dummy variable*. If you systematically replace it with another variable, the resulting integral will still be the same. For instance,

$$\int_0^1 x^2 dx = \left[\frac{1}{3}x^3\right]_{x=0}^1 = \frac{1}{3},$$

and if you replace x by φ you still get

$$\int_0^1 \varphi^2 d\varphi = \left[\frac{1}{3}\varphi^3\right]_{\varphi=0}^1 = \frac{1}{3}.$$

Another way to appreciate that the integration variable is a dummy variable is to look at the Fundamental Theorem again:

$$\int_a^b f(x) dx = F(b) - F(a).$$

The right hand side tells you that the value of the integral depends on a and b , and has absolutely nothing to do with the variable x .

4. Exercises

330. What is a Riemann sum of a function $y = f(x)$?

331. Let f be the function $f(x) = 1 - x^2$.

Draw the graph of $f(x)$ with $0 \leq x \leq 2$.

Compute the Riemann-sum for the partition

$$0 < \frac{1}{3} < 1 < \frac{3}{2} < 2$$

of the interval $[a, b] = [0, 2]$ if you choose each c_k to be the left endpoint of the interval it belongs to. Draw the corresponding rectangles (add them to your drawing of the graph of f).

Then compute the Riemann-sum you get if you choose the c_k to be the right endpoint of the interval it belongs to. Make a new drawing of the graph of f and include the rectangles corresponding to the right endpoint Riemann-sum.

332. Group Problem.

Look at figure 1 (top). Which choice of intermediate points c_1, \dots, c_6 leads to the smallest Riemann sum? Which choice would give you the largest Riemann-sum?

(Note: in this problem you're not allowed to change the division points x_i , only the points c_i in between them.)

Find an antiderivative $F(x)$ for each of the following functions $f(x)$. Finding antiderivatives involves a fair amount of guess work, but with experience it gets easier to guess antiderivatives.

333. $f(x) = 2x + 1$

334. $f(x) = 1 - 3x$

335. $f(x) = x^2 - x + 11$

336. $f(x) = x^4 - x^2$

337. $f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4}$

338. $f(x) = \frac{1}{x}$

339. $f(x) = e^x$

340. $f(x) = \frac{2}{x}$

341. $f(x) = e^{2x}$

342. $f(x) = \frac{1}{2+x}$

343. $f(x) = \frac{e^x - e^{-x}}{2}$

344. $f(x) = \frac{1}{1+x^2}$

345. $f(x) = \frac{e^x + e^{-x}}{2}$

346. $f(x) = \frac{1}{\sqrt{1-x^2}}$

347. $f(x) = \sin x$

348. $f(x) = \frac{2}{1-x}$

349. $f(x) = \cos x$

350. $f(x) = \cos 2x$

351. $f(x) = \sin(x - \pi/3)$

352. $f(x) = \sin x + \sin 2x$

353. $f(x) = 2x(1 + x^2)^5$

In each of the following exercises you should compute the area of the indicated region, and also of the smallest enclosing rectangle with horizontal and vertical sides.

Before computing anything draw the region.

354. The region between the vertical lines $x = 0$ and $x = 1$, and between the x -axis and the graph of $y = x^3$.

355. The region between the vertical lines $x = 0$ and $x = 1$, and between the x -axis and the graph of $y = x^n$ (here $n > 0$, draw for $n = \frac{1}{2}, 1, 2, 3, 4$).

356. The region above the graph of $y = \sqrt{x}$, below the line $y = 2$, and between the vertical lines $x = 0$, $x = 4$.

357. The region above the x -axis and below the graph of $f(x) = x^2 - x^3$.

358. The region above the x -axis and below the graph of $f(x) = 4x^2 - x^4$.

359. The region above the x -axis and below the graph of $f(x) = 1 - x^4$.

360. The region above the x -axis, below the graph of $f(x) = \sin x$, and between $x = 0$ and $x = \pi$.

361. The region above the x -axis, below the graph of $f(x) = 1/(1 + x^2)$ (a curve known as *Maria Agnesi's witch*), and between $x = 0$ and $x = 1$.

362. The region between the graph of $y = 1/x$ and the x -axis, and between $x = a$ and $x = b$ (here $0 < a < b$ are constants, e.g. choose $a = 1$ and $b = \sqrt{2}$ if you have something against either letter a or b .)

363. The region above the x -axis and below the graph of

$$f(x) = \frac{1}{1+x} + \frac{x}{2} - 1.$$

364. Compute

$$\int_0^1 \sqrt{1-x^2} dx$$

without finding an antiderivative for $\sqrt{1-x^2}$ (you can find such an antiderivative, but it's not easy. This integral is the area of some region: which region is it, and what is that area?)

365. **Group Problem.**

Compute these integrals without finding antiderivatives.

$$I = \int_0^{1/2} \sqrt{1-x^2} dx$$

$$J = \int_{-1}^1 |1-x| dx$$

$$K = \int_{-1}^1 |2-x| dx$$

5. The indefinite integral

The fundamental theorem tells us that in order to compute the integral of some function f over an interval $[a, b]$ you should first find an antiderivative F of f . In practice, much of the effort required to find an integral goes into finding the antiderivative. In order to simplify the computation of the integral

$$(49) \quad \int_a^b f(x) dx = F(b) - F(a)$$

the following notation is commonly used for the antiderivative:

$$(50) \quad F(x) = \int f(x) dx.$$

For instance,

$$\int x^2 dx = \frac{1}{3}x^3, \quad \int \sin 5x dx = -\frac{1}{5}\cos 5x, \quad \text{etc} \dots$$

The integral which appears here does not have the integration bounds a and b . It is called an *indefinite integral*, as opposed to the integral in (49) which is called a *definite integral*. You use the indefinite integral if you expect the computation of the antiderivative to be a lengthy affair, and you do not want to write the integration bounds a and b all the time.

It is important to distinguish between the two kinds of integrals. Here is a list of differences:

| INDEFINITE INTEGRAL | DEFINITE INTEGRAL |
|--|--|
| $\int f(x)dx$ is a function of x . | $\int_a^b f(x)dx$ is a number. |
| By definition $\int f(x)dx$ is <i>any function</i> of x whose derivative is $f(x)$. | $\int_a^b f(x)dx$ was defined in terms of Riemann sums and can be interpreted as “area under the graph of $y = f(x)$ ”, at least when $f(x) > 0$. |
| x is not a dummy variable, for example, $\int 2x dx = x^2 + C$ and $\int 2t dt = t^2 + C$ are functions of different variables, so they are not equal. | x is a dummy variable, for example, $\int_0^1 2x dx = 1$, and $\int_0^1 2t dt = 1$, so $\int_0^1 2x dx = \int_0^1 2t dt$. |

5.1. You can always check the answer. Suppose you want to find an antiderivative of a given function $f(x)$ and after a long and messy computation which you don't really trust you get an “answer”, $F(x)$. You can then throw away the dubious computation and differentiate the $F(x)$ you had found. If $F'(x)$ turns out to be equal to $f(x)$, then your $F(x)$ is indeed an antiderivative and your computation isn't important anymore.

For example, suppose that we want to find $\int \ln x dx$. My cousin Louie says it might be $F(x) = x \ln x - x$. Let's see if he's right:

$$\frac{d}{dx}(x \ln x - x) = x \cdot \frac{1}{x} + 1 \cdot \ln x - 1 = \ln x.$$

Who knows how Louie thought of this¹, but it doesn't matter: he's right! We now know that $\int \ln x dx = x \ln x - x + C$.

5.2. About “+C”. Let $f(x)$ be a function defined on some interval $a \leq x \leq b$. If $F(x)$ is an antiderivative of $f(x)$ on this interval, then for any constant C the function $\tilde{F}(x) = F(x) + C$ will also be an antiderivative of $f(x)$. So one given function $f(x)$ has many different antiderivatives, obtained by adding different constants to one given antiderivative.

5.3. Theorem. If $F_1(x)$ and $F_2(x)$ are antiderivatives of the same function $f(x)$ on some interval $a \leq x \leq b$, then there is a constant C such that $F_1(x) = F_2(x) + C$.

PROOF. Consider the difference $G(x) = F_1(x) - F_2(x)$. Then $G'(x) = F_1'(x) - F_2'(x) = f(x) - f(x) = 0$, so that $G(x)$ must be constant. Hence $F_1(x) - F_2(x) = C$ for some constant. \square

It follows that there is some ambiguity in the notation $\int f(x) dx$. Two functions $F_1(x)$ and $F_2(x)$ can both equal $\int f(x) dx$ without equaling each other. When this happens, they (F_1 and F_2) differ by a constant. This can sometimes lead to confusing situations, e.g. you can check that

$$\begin{aligned} \int 2 \sin x \cos x dx &= \sin^2 x \\ \int 2 \sin x \cos x dx &= -\cos^2 x \end{aligned}$$

are both correct. (Just differentiate the two functions $\sin^2 x$ and $-\cos^2 x$!) These two answers look different until you realize that because of the trig identity $\sin^2 x + \cos^2 x = 1$ they really only differ by a constant: $\sin^2 x = -\cos^2 x + 1$.

To avoid this kind of confusion we will from now on never forget to include the “arbitrary constant +C” in our answer when we compute an antiderivative.

$$\begin{aligned}
\int f(x) dx &= F(x) + C \\
\int x^n dx &= \frac{x^{n+1}}{n+1} + C && \text{for all } n \neq -1 \\
\int \frac{1}{x} dx &= \ln |x| + C && \text{(Note the absolute values)} \\
\int e^x dx &= e^x + C \\
\int a^x dx &= \frac{a^x}{\ln a} + C && \text{(don't memorize: use } a^x = e^{x \ln a} \text{)} \\
\int \sin x dx &= -\cos x + C \\
\int \cos x dx &= \sin x + C \\
\int \tan x dx &= -\ln |\cos x| + C && \text{(Note the absolute values)} \\
\int \frac{1}{1+x^2} dx &= \arctan x + C \\
\int \frac{1}{\sqrt{1-x^2}} dx &= \arcsin x + C
\end{aligned}$$

The following integral is also useful, but not as important as the ones above:

$$\int \frac{dx}{\cos x} = \frac{1}{2} \ln \frac{1 + \sin x}{1 - \sin x} + C \text{ for } -\frac{\pi}{2} < x < \frac{\pi}{2}.$$

Table 1. The list of the standard integrals everyone should know

Table 1 lists a number of antiderivatives which you should know. All of these integrals should be familiar from the differentiation rules we have learned so far, except for the integrals of $\tan x$ and of $\frac{1}{\cos x}$. You can check those by differentiation (using $\ln \frac{a}{b} = \ln a - \ln b$ simplifies things a bit).

6. Properties of the Integral

Just as we had a list of properties for the limits and derivatives of sums and products of functions, the integral has similar properties.

Suppose we have two functions $f(x)$ and $g(x)$ with antiderivatives $F(x)$ and $G(x)$, respectively. Then we know that

$$\frac{d}{dx} \{F(x) + G(x)\} = F'(x) + G'(x) = f(x) + g(x),$$

in words, $F + G$ is an antiderivative of $f + g$, which we can write as

$$(51) \quad \int \{f(x) + g(x)\} dx = \int f(x) dx + \int g(x) dx.$$

Similarly, $\frac{d}{dx}(cF(x)) = cF'(x) = cf(x)$ implies that

$$(52) \quad \int cf(x) dx = c \int f(x) dx$$

if c is a constant.

¹He took math 222 and learned to integrate by parts.

These properties imply analogous properties for the definite integral. For any pair of functions on an interval $[a, b]$ one has

$$(53) \quad \int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx,$$

and for any function f and constant c one has

$$(54) \quad \int_a^b cf(x) dx = c \int_a^b f(x) dx.$$

Definite integrals have one other property for which there is no analog in indefinite integrals: if you split the interval of integration into two parts, then the integral over the whole is the sum of the integrals over the parts. The following theorem says it more precisely.

6.1. Theorem. *Given $a < c < b$, and a function on the interval $[a, b]$ then*

$$(55) \quad \int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

PROOF. Let F be an antiderivative of f . Then

$$\int_a^c f(x) dx = F(c) - F(a) \text{ and } \int_c^b f(x) dx = F(b) - F(c),$$

so that

$$\begin{aligned} \int_a^b f(x) dx &= F(b) - F(a) \\ &= F(b) - F(c) + F(c) - F(a) \\ &= \int_a^c f(x) dx + \int_c^b f(x) dx. \end{aligned}$$

□

So far we have always assumed that $a < b$ in all indefinite integrals $\int_a^b \dots$. The fundamental theorem suggests that when $b < a$, we should define the integral as

$$(56) \quad \int_a^b f(x) dx = F(b) - F(a) = -(F(a) - F(b)) = - \int_b^a f(x) dx.$$

For instance,

$$\int_1^0 x dx = - \int_0^1 x dx = -\frac{1}{2}.$$

7. The definite integral as a function of its integration bounds

Consider the expression

$$I = \int_0^x t^2 dt.$$

What does I depend on? To see this, you calculate the integral and you find

$$I = \left[\frac{1}{3} t^3 \right]_0^x = \frac{1}{3} x^3 - \frac{1}{3} 0^3 = \frac{1}{3} x^3.$$

So the integral depends on x . It does not depend on t , since t is a “dummy variable” (see §3.3 where we already discussed this point.)

In this way you can use integrals to define new functions. For instance, we could define

$$I(x) = \int_0^x t^2 dt,$$

which would be a roundabout way of defining the function $I(x) = x^3/3$. Again, since t is a dummy variable we can replace it by any other variable we like. Thus

$$I(x) = \int_0^x \alpha^2 d\alpha$$

defines the same function (namely, $I(x) = \frac{1}{3}x^3$).

The previous example does not define a new function ($I(x) = x^3/3$). An example of a *new* function defined by an integral is the “error-function” from statistics. It is given by

$$(57) \quad \operatorname{erf}(x) \stackrel{\text{def}}{=} \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

so $\operatorname{erf}(x)$ is the area of the shaded region in figure 3. The integral in (57) cannot be computed in terms of the

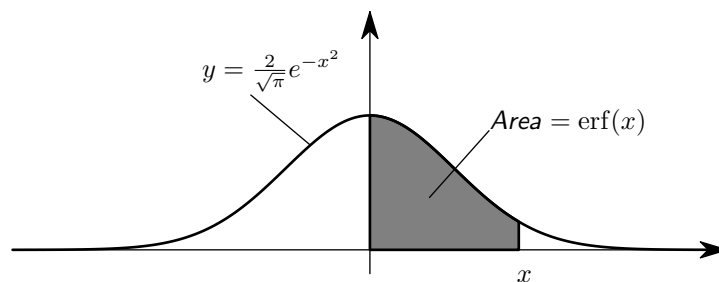


Figure 3. Definition of the Error function.

standard functions (square and higher roots, sine, cosine, exponential and logarithms). Since the integral in (57) occurs very often in statistics (in relation with the so-called normal distribution) it has been given a name, namely, “ $\operatorname{erf}(x)$ ”.

How do you differentiate a function that is defined by an integral? The answer is simple, for if $f(x) = F'(x)$ then the fundamental theorem says that

$$\int_a^x f(t) dt = F(x) - F(a),$$

and therefore

$$\frac{d}{dx} \int_a^x f(t) dt = \frac{d}{dx} \{F(x) - F(a)\} = F'(x) = f(x),$$

i.e.

$$\frac{d}{dx} \int_a^x f(t) dt = f(x).$$

A similar calculation gives you

$$\frac{d}{dx} \int_x^b f(t) dt = -f(x).$$

So what is the derivative of the error function? We have

$$\begin{aligned} \operatorname{erf}'(x) &= \frac{d}{dx} \left\{ \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \right\} \\ &= \frac{2}{\sqrt{\pi}} \frac{d}{dx} \int_0^x e^{-t^2} dt \\ &= \frac{2}{\sqrt{\pi}} e^{-x^2}. \end{aligned}$$

8. Method of substitution

The chain rule says that

$$\frac{dF(G(x))}{dx} = F'(G(x)) \cdot G'(x),$$

so that

$$\int F'(G(x)) \cdot G'(x) dx = F(G(x)) + C.$$

8.1. Example. Consider the function $f(x) = 2x \sin(x^2 + 3)$. It does not appear in the list of standard antiderivatives we know by heart. But we do notice² that $2x = \frac{d}{dx}(x^2 + 3)$. So let's call $G(x) = x^2 + 3$, and $F(u) = -\cos u$, then

$$F(G(x)) = -\cos(x^2 + 3)$$

and

$$\frac{dF(G(x))}{dx} = \underbrace{\sin(x^2 + 3)}_{F'(G(x))} \cdot \underbrace{2x}_{G'(x)} = f(x),$$

so that

$$(58) \quad \int 2x \sin(x^2 + 3) dx = -\cos(x^2 + 3) + C.$$

8.2. Leibniz' notation for substitution. The most transparent way of computing an integral by substitution is by following Leibniz and introduce new variables. Thus to do the integral

$$\int f(G(x))G'(x) dx$$

where $f(u) = F'(u)$, we introduce the substitution $u = G(x)$, and agree to write

$$du = dG(x) = G'(x) dx.$$

Then we get

$$\int f(G(x))G'(x) dx = \int f(u) du = F(u) + C.$$

At the end of the integration we must remember that u really stands for $G(x)$, so that

$$\int f(G(x))G'(x) dx = F(u) + C = F(G(x)) + C.$$

As an example, let's do the integral (58) using Leibniz' notation. We want to find

$$\int 2x \sin(x^2 + 3) dx$$

and decide to substitute $z = x^2 + 3$ (the substitution variable doesn't always have to be called u). Then we compute

$$dz = d(x^2 + 3) = 2x dx \text{ and } \sin(x^2 + 3) = \sin z,$$

so that

$$\int 2x \sin(x^2 + 3) dx = \int \sin z dz = -\cos z + C.$$

Finally we get rid of the substitution variable z , and we find

$$\int 2x \sin(x^2 + 3) dx = -\cos(x^2 + 3) + C.$$

When we do integrals in this calculus class, we always get rid of the substitution variable because it is a variable we invented, and which does not appear in the original problem. But if you are doing an integral which appears in some longer discussion of a real-life (or real-lab) situation, then it may be that the substitution variable actually has a meaning (e.g. "the effective stoichiometric modality of CQF self-inhibition") in which

² You *will* start noticing things like this after doing several examples.

case you may want to skip the last step and leave the integral in terms of the (meaningful) substitution variable.

8.3. Substitution for definite integrals. For definite integrals the chain rule

$$\frac{d}{dx}(F(G(x))) = F'(G(x))G'(x) = f(G(x))G'(x)$$

implies

$$\int_a^b f(G(x))G'(x) dx = F(G(b)) - F(G(a)).$$

which you can also write as

$$(59) \quad \int_{x=a}^b f(G(x))G'(x) dx = \int_{u=G(a)}^{G(b)} f(u) du.$$

8.4. Example of substitution in a definite integral. Let's compute

$$\int_0^1 \frac{x}{1+x^2} dx,$$

using the substitution $u = G(x) = 1 + x^2$. Since $du = 2x dx$, the associated *indefinite* integral is

$$\int \underbrace{\frac{1}{1+x^2}}_{\frac{1}{u}} \underbrace{x dx}_{\frac{1}{2} du} = \frac{1}{2} \int \frac{1}{u} du.$$

To find the definite integral you must compute the new integration bounds $G(0)$ and $G(1)$ (see equation (59).) If x runs between $x = 0$ and $x = 1$, then $u = G(x) = 1 + x^2$ runs between $u = 1 + 0^2 = 1$ and $u = 1 + 1^2 = 2$, so the definite integral we must compute is

$$(60) \quad \int_0^1 \frac{x}{1+x^2} dx = \frac{1}{2} \int_1^2 \frac{1}{u} du,$$

which is in our list of memorable integrals. So we find

$$\int_0^1 \frac{x}{1+x^2} dx = \frac{1}{2} \int_1^2 \frac{1}{u} du = \frac{1}{2} [\ln u]_1^2 = \frac{1}{2} \ln 2.$$

Sometimes the integrals in (60) are written as

$$\int_{x=0}^1 \frac{x}{1+x^2} dx = \frac{1}{2} \int_{u=1}^2 \frac{1}{u} du,$$

to emphasize (and remind yourself) to which variable the bounds in the integral refer.

9. Exercises

Compute these derivatives:

$$366. \quad \frac{d}{dx} \int_0^x (1+t^2)^4 dt$$

$$367. \quad \frac{d}{dx} \int_x^1 \ln z dz$$

$$368. \quad \frac{d}{dt} \int_0^t \frac{dx}{1+x^2}$$

$$369. \quad \frac{d}{dt} \int_0^{1/t} \frac{dx}{1+x^2}$$

$$370. \quad \frac{d}{dx} \int_x^{2x} s^2 ds$$

$$371. \quad \frac{d}{dq} \int_{-q}^q \frac{dx}{1-x^2} \quad [\text{Which values of } q \text{ are allowed here?}]$$

$$372. \quad \frac{d}{dt} \int_0^{t^2} e^{2x} dx$$

373. Group Problem.

You can see the graph of the error function at http://en.wikipedia.org/wiki/Error_function

(a) Compute the second derivative of the error function. How many inflection points does the graph of the error function have?

(b) The graph of the error function on Wikipedia shows that $\operatorname{erf}(x)$ is negative when $x < 0$. But the error

function is defined as an integral of a positive function so it should be positive. Is Wikipedia wrong? Explain.

Compute the following indefinite integrals:

$$374. \int \{6x^5 - 2x^{-4} - 7x\} dx$$

$$375. \int \{+3/x - 5 + 4e^x + 7^x\} dx$$

$$376. \int (x/a + a/x + x^a + a^x + ax) dx$$

$$377. \int \{\sqrt{x} - \sqrt[3]{x^4} + \frac{7}{\sqrt[3]{x^2}} - 6e^x + 1\} dx$$

$$378. \int \{2^x + (\frac{1}{2})^x\} dx$$

$$379. \int_{-2}^4 (3x - 5) dx$$

$$380. \int_1^4 x^{-2} dx \quad (\text{hm} \dots)$$

$$381. \int_1^4 t^{-2} dt \quad (!)$$

$$382. \int_1^4 x^{-2} dt \quad (!!!)$$

$$383. \int_0^1 (1 - 2x - 3x^2) dx$$

$$384. \int_1^2 (5x^2 - 4x + 3) dx$$

$$385. \int_{-3}^0 (5y^4 - 6y^2 + 14) dy$$

$$386. \int_0^1 (y^9 - 2y^5 + 3y) dy$$

$$387. \int_0^4 \sqrt{x} dx$$

$$388. \int_0^1 x^{3/7} dx$$

$$389. \int_1^3 \left(\frac{1}{t^2} - \frac{1}{t^4} \right) dt$$

$$390. \int_1^2 \frac{t^6 - t^2}{t^4} dt$$

$$391. \int_1^2 \frac{x^2 + 1}{\sqrt{x}} dx$$

$$392. \int_0^2 (x^3 - 1)^2 dx$$

$$393. \int_0^1 u(\sqrt{u} + \sqrt[3]{u}) du$$

$$394. \int_1^2 (x + 1/x)^2 dx$$

$$395. \int_3^3 \sqrt{x^5 + 2} dx$$

$$396. \int_1^{-1} (x - 1)(3x + 2) dx$$

$$397. \int_1^4 (\sqrt{t} - 2/\sqrt{t}) dt$$

$$398. \int_1^8 \left(\sqrt[3]{r} + \frac{1}{\sqrt[3]{r}} \right) dr$$

$$399. \int_{-1}^0 (x + 1)^3 dx$$

$$400. \int_{-5}^{-2} \frac{x^4 - 1}{x^2 + 1} dx$$

$$401. \int_1^e \frac{x^2 + x + 1}{x} dx$$

$$402. \int_4^9 \left(\sqrt{x} + \frac{1}{\sqrt{x}} \right)^2 dx$$

$$403. \int_0^1 \left(\sqrt[4]{x^5} + \sqrt[5]{x^4} \right) dx$$

$$404. \int_1^8 \frac{x - 1}{\sqrt[3]{x^2}} dx$$

$$405. \int_{\pi/4}^{\pi/3} \sin t dt$$

$$406. \int_0^{\pi/2} (\cos \theta + 2 \sin \theta) d\theta$$

$$407. \int_0^{\pi/2} (\cos \theta + \sin 2\theta) d\theta$$

$$408. \int_{2\pi/3}^{\pi} \frac{\tan x}{\cos x} dx$$

$$409. \int_{\pi/3}^{\pi/2} \frac{\cot x}{\sin x} dx$$

$$410. \int_1^{\sqrt{3}} \frac{6}{1 + x^2} dx$$

$$411. \int_0^{0.5} \frac{dx}{\sqrt{1 - x^2}}$$

$$412. \int_4^8 (1/x) dx$$

$$413. \int_{\ln 3}^{\ln 6} 8e^x dx$$

$$414. \int_8^9 2^t dt$$

$$415. \int_{-e^2}^{-e} \frac{3}{x} dx$$

$$416. \int_{-2}^3 |x^2 - 1| dx$$

417. $\int_{-1}^2 |x - x^2| dx$

418. $\int_{-1}^2 (x - 2|x|) dx$

419. $\int_0^2 (x^2 - |x - 1|) dx$

420. $\int_0^2 f(x) dx$ where $f(x) = \begin{cases} x^4, & \text{if } 0 \leq x < 1, \\ x^5, & \text{if } 1 \leq x \leq 2. \end{cases}$

421. $\int_{-\pi}^{\pi} f(x) dx$ where $f(x) = \begin{cases} x, & \text{if } -\pi \leq x \leq 0, \\ \sin x, & \text{if } 0 < x \leq \pi. \end{cases}$

422. Compute

$$I = \int_0^2 2x(1 + x^2)^3 dx$$

in two different ways:

(a) Expand $(1 + x^2)^3$, multiply with $2x$, and integrate each term.

(b) Use the substitution $u = 1 + x^2$.

423. Compute

$$I_n = \int 2x(1 + x^2)^n dx.$$

424. If $f'(x) = x - 1/x^2$ and $f(1) = 1/2$ find $f(x)$.

425. Sketch the graph of the curve $y = \sqrt{x+1}$ and determine the area of the region enclosed by the curve, the x -axis and the lines $x = 0$, $x = 4$.

426. Find the area under the curve $y = \sqrt{6x+4}$ and above the x -axis between $x = 0$ and $x = 2$. Draw a sketch of the curve.

427. Graph the curve $y = 2\sqrt{1-x^2}$, $x \in [0, 1]$, and find the area enclosed between the curve and the x -axis. (Don't evaluate the integral, but compare with the area under the graph of $y = \sqrt{1-x^2}$.)

428. Determine the area under the curve $y = \sqrt{a^2 - x^2}$ and between the lines $x = 0$ and $x = a$.

429. Graph the curve $y = 2\sqrt{9-x^2}$ and determine the area enclosed between the curve and the x -axis.

430. Graph the area between the curve $y^2 = 4x$ and the line $x = 3$. Find the area of this region.

431. Find the area bounded by the curve $y = 4 - x^2$ and the lines $y = 0$ and $y = 3$.

432. Find the area enclosed between the curve $y = \sin 2x$, $0 \leq x \leq \pi/4$ and the axes.

433. Find the area enclosed between the curve $y = \cos 2x$, $0 \leq x \leq \pi/4$ and the axes.

434. Graph $y^2 + 1 = x$, and find the area enclosed by the curve and the line $x = 2$.

435. Find the area of the region bounded by the parabola $y^2 = 4x$ and the line $y = 2x$.

436. Find the area bounded by the curve $y = x(2-x)$ and the line $x = 2y$.

437. Find the area bounded by the curve $x^2 = 4y$ and the line $x = 4y - 2$.

438. Calculate the area of the region bounded by the parabolas $y = x^2$ and $x = y^2$.

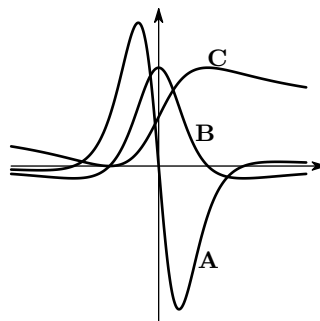
439. Find the area of the region included between the parabola $y^2 = x$ and the line $x + y = 2$.

440. Find the area of the region bounded by the curves $y = \sqrt{x}$ and $y = x$.

441. Group Problem.

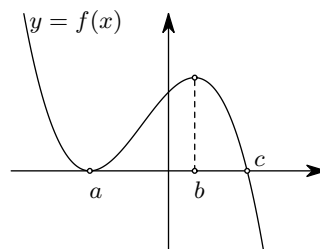
You asked your assistant Joe to produce graphs of a function $f(x)$, its derivative $f'(x)$ and an antiderivative $F(x)$ of $f(x)$.

Unfortunately Joe simply labelled the graphs "A," "B," and "C," and now he doesn't remember which graph is f , which is f' and which is F . Identify which graph is which and explain your answer.



442. Group Problem.

Below is the graph of a function $y = f(x)$.



The function $F(x)$ (graph not shown) is an antiderivative of $f(x)$. Which among the following statements true?

(a) $F(a) = F(c)$

(b) $F(b) = 0$

(c) $F(b) > F(c)$

(d) The graph of $y = F(x)$ has **two** inflection points?

Use a substitution to evaluate the following integrals.

$$443. \int_1^2 \frac{u \, du}{1+u^2}$$

$$444. \int_0^5 \frac{x \, dx}{\sqrt{x+1}}$$

$$445. \int_1^2 \frac{x^2 \, dx}{\sqrt{2x+1}}$$

$$446. \int_0^5 \frac{s \, ds}{\sqrt[3]{s+2}}$$

$$447. \int_1^2 \frac{x \, dx}{1+x^2}$$

$$448. \int_0^\pi \cos(\theta + \frac{\pi}{3}) d\theta$$

$$449. \int \sin \frac{\pi+x}{5} dx$$

$$450. \int \frac{\sin 2x}{\sqrt{1+\cos 2x}} dx$$

$$451. \int_{\pi/4}^{\pi/3} \sin^2 \theta \cos \theta \, d\theta$$

$$452. \int_2^3 \frac{1}{r \ln r}, dr$$

$$453. \int \frac{\sin 2x}{1+\cos^2 x} dx$$

$$454. \int \frac{\sin 2x}{1+\sin x} dx$$

$$455. \int_0^1 z \sqrt{1-z^2} dz$$

$$456. \int_1^2 \frac{\ln 2x}{x} dx$$

$$457. \int_{\xi=0}^{\sqrt{2}} \xi(1+2\xi^2)^{10} d\xi$$

$$458. \int_2^3 \sin \rho (\cos 2\rho)^4 d\rho$$

$$459. \int \alpha e^{-\alpha^2} d\alpha$$

$$460. \int \frac{e^{\frac{1}{t}}}{t^2} dt$$

Applications of the integral

The integral appears as the answer to many different questions. In this chapter we will describe a number of “things which are an integral.” In each example there is a quantity we want to compute, and which we can approximate through Riemann-sums. After letting the partition become arbitrarily fine we then find that the quantity we are looking for is given by an integral. The derivations are an important part of the subject.

1. Areas between graphs

Suppose you have two functions f and g on an interval $[a, b]$, one of which is always larger than the other, i.e. for which you know that $f(x) \leq g(x)$ for all x in the interval $[a, b]$. Then the area of the region between the graphs of the two functions is

$$(61) \quad \text{Area} = \int_a^b (g(x) - f(x)) dx.$$

To get this formula you approximate the region by a large number of thin rectangles. Choose a partition $a = x_0 < x_1 < \cdots < x_n = b$ of the interval $[a, b]$; choose a number c_k in each interval $[x_{k-1}, x_k]$; form the rectangles

$$x_{k-1} \leq x \leq x_k, \quad f(c_k) \leq y \leq g(c_k).$$

The area of this rectangle is

$$\text{width} \times \text{height} = \Delta x_k \times (g(c_k) - f(c_k)).$$

Hence the combined area of the rectangles is

$$R = (g(c_1) - f(c_1))\Delta x_1 + \cdots + (g(c_n) - f(c_n))\Delta x_n$$

which is just the Riemann-sum for the integral

$$I = \int_a^b (g(x) - f(x)) dx.$$

So,

- (1) since the area of the region between the graphs of f and g is the limit of the combined areas of the rectangles,

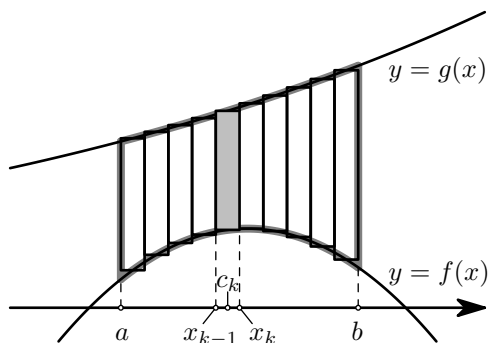


Figure 1. Finding the area between two graphs using Riemann-sums.

- (2) and since this combined area is equal to the Riemann sum R ,
 (3) and since the Riemann-sums R converge to the integral I ,

we conclude that the area between the graphs of f and g is exactly the integral I .

2. Exercises

- 461.** Find the area of the region bounded by the parabola $y^2 = 4x$ and the line $y = 2x$.
462. Find the area bounded by the curve $y = x(2 - x)$ and the line $x = 2y$.
463. Find the area bounded by the curve $x^2 = 4y$ and the line $x = 4y - 2$.
464. Calculate the area of the region bounded by the parabolas $y = x^2$ and $x = y^2$.
465. Find the area of the region included between the parabola $y^2 = x$ and the line $x + y = 2$.
466. Find the area of the region bounded by the curves $y = \sqrt{x}$ and $y = x$.
467. Use integration to find the area of the triangular region bounded by the lines $y = 2x + 1$, $y = 3x + 1$ and $x = 4$.
468. Find the area bounded by the parabola $x^2 - 2 = y$ and the line $x + y = 0$.
469. Where do the graphs of $f(x) = x^2$ and $g(x) = 3/(2 + x^2)$ intersect? Find the area of the region which lies above the graph of g and below the graph of f .

(Hint: if you need to integrate $1/(2 + x^2)$ you could substitute $x = u\sqrt{2}$.)

- 470.** Graph the curve $y = (1/2)x^2 + 1$ and the straight line $y = x + 1$ and find the area between the curve and the line.
471. Find the area of the region between the parabolas $y^2 = x$ and $x^2 = 16y$.
472. Find the area of the region enclosed by the parabola $y^2 = 4ax$ and the line $y = mx$.
473. Find a so that the curves $y = x^2$ and $y = a \cos x$ intersect at the points $(x, y) = (\frac{\pi}{4}, \frac{\pi^2}{16})$. Then find the area between these curves.
474. Group Problem.

Write a definite integral whose value is the area of the region between the two circles $x^2 + y^2 = 1$ and $(x - 1)^2 + y^2 = 1$. Find this area. If you cannot evaluate the integral by calculus you may use geometry to find the area. Hint: The part of a circle cut off by a line is a circular sector with a triangle removed.

3. Cavalieri's principle and volumes of solids

You can use integration to derive the formulas for volumes of spheres, cylinder, cones, and many many more solid objects in a systematic way. In this section we'll see the "method of slicing."

3.1. Example – Volume of a pyramid. As an example let's compute the volume of a pyramid whose base is a square of side 1, and whose height is 1. Our strategy will be to divide the pyramid into thin horizontal slices whose volumes we can compute, and to add the volumes of the slices to get the volume of the pyramid.

To construct the slices we choose a partition of the (height) interval $[0, 1]$ into N subintervals, i.e. we pick numbers

$$0 = x_0 < x_1 < x_2 < \cdots < x_{N-1} < x_N = 1,$$

and as usual we set $\Delta x_k = x_k - x_{k-1}$, we define the mesh size of the partition to be the largest of the Δx_k .

The k^{th} slice consists of those points on the pyramid whose height is between x_{k-1} and x_k . The intersection of the pyramid with the plane at height x is a square, and by similarity the length of the side of this square is $1 - x$. Therefore the bottom of the k^{th} slice is a square with side $1 - x_{k-1}$, and its top is a square with side $1 - x_k$. The height of the slice is $x_k - x_{k-1} = \Delta x_k$.

Thus the k^{th} slice *contains* a block of height Δx_k whose base is a square with sides $1 - x_k$, and its volume must therefore be larger than $(1 - x_k)^2 \Delta x_k$. On the other hand the k^{th} slice *is contained in* a block of the same height whose base is a square with sides $1 - x_{k-1}$. The volume of the slice is therefore not more than $(1 - x_{k-1})^2 \Delta x_k$. So we have

$$(1 - x_k)^2 \Delta x_k < \text{volume of } k^{\text{th}} \text{ slice} < (1 - x_{k-1})^2 \Delta x_k.$$

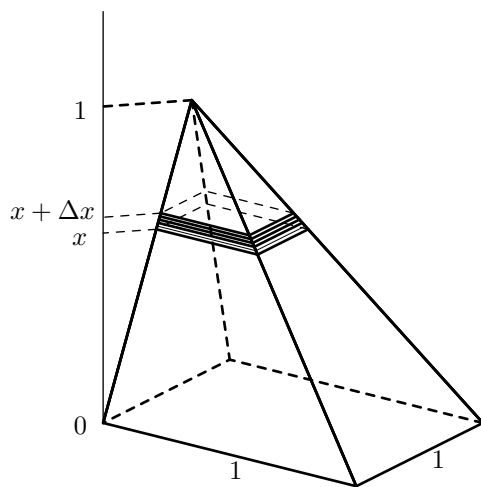


Figure 2. The slice at height x is a square with side $1 - x$.

Therefore there is some c_k in the interval $[x_{k-1}, x_k]$ such that

$$\text{volume of } k^{\text{th}} \text{ slice} = (1 - c_k)^2 \Delta x_k.$$

Adding the volumes of the slices we find that the volume V of the pyramid is given by

$$V = (1 - c_1)^2 \Delta x_1 + \cdots + (1 - c_N)^2 \Delta x_N.$$

The right hand side in this equation is a Riemann sum for the integral

$$I = \int_0^1 (1 - x)^2 dx$$

and therefore we have

$$I = \lim_{N \rightarrow \infty} \{(1 - c_1)^2 \Delta x_1 + \cdots + (1 - c_N)^2 \Delta x_N\} = V.$$

Compute the integral and you find that the volume of the pyramid is

$$V = \frac{1}{3}.$$

3.2. General case. The “method of slicing” which we just used to compute the volume of a pyramid works for solids of any shape. The strategy always consists of dividing the solid into many thin (horizontal) slices, compute their volumes, and recognize that the total volume of the slices is a Riemann sum for some integral. That integral then is the volume of the solid.

To be more precise, let a and b be the heights of the lowest and highest points on the solid, and let $a = x_0 < x_1 < x_2 < \cdots < x_{N-1} < x_N = b$ be a partition of the interval $[a, b]$. Such a partition divides the solid into N distinct slices, where slice number k consists of all points in the solid whose height is between x_{k-1} and x_k . The thickness of the k^{th} slice is $\Delta x_k = x_k - x_{k-1}$. If

$$A(x) = \text{area of the intersection of the solid with the plane at height } x.$$

then we can approximate the volume of the k^{th} slice by

$$A(c_k) \Delta x_k$$

where c_k is any number (height) between x_{k-1} and x_k .

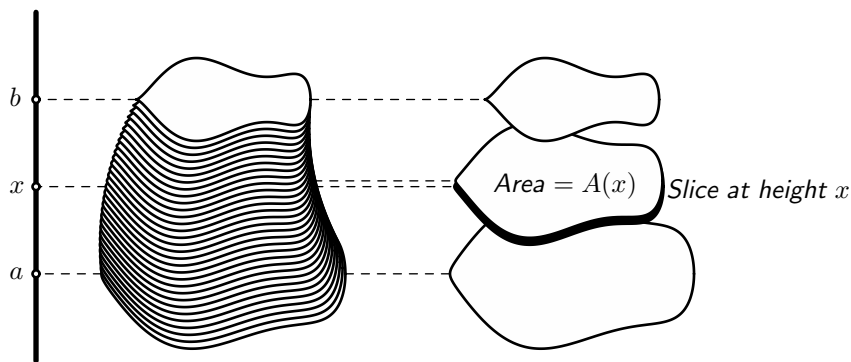


Figure 3. Slicing a solid to compute its volume. The volume of one slice is approximately the product of its thickness (Δx) and the area $A(x)$ of its top. Summing the volume $A(x)\Delta x$ over all slices leads approximately to the integral $\int_a^b f(x)dx$.

The total volume of all slices is therefore approximately

$$V \approx A(c_1)\Delta x_1 + \cdots + A(c_N)\Delta x_N.$$

While this formula only holds approximately, we expect the approximation to get better as we make the partition finer, and thus

$$(62) \quad V = \lim_{N \rightarrow \infty} \{A(c_1)\Delta x_1 + \cdots + A(c_N)\Delta x_N\}.$$

On the other hand the sum on the right is a Riemann sum for the integral $I = \int_a^b A(x)dx$, so the limit is exactly this integral. Therefore we have

$$(63) \quad V = \int_a^b A(x)dx.$$

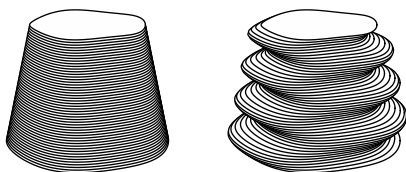


Figure 4. Cavalieri's principle. Both solids consist of a pile of horizontal slices. The solid on the right was obtained from the solid on the left by sliding some of the slices to the left and others to the right. This operation does not affect the volumes of the slices, and hence both solids have the same volume.

3.3. Cavalieri's principle. The formula (63) for the volume of a solid which we have just derived shows that the volume only depends on the areas $A(x)$ of the cross sections of the solid, and not on the particular shape these cross sections may have. This observation is older than calculus itself and goes back at least to Bonaventura Cavalieri (1598 – 1647) who said: *If the intersections of two solids with a horizontal plane always have the same area, no matter what the height of the horizontal plane may be, then the two solids have the same volume.*

This principle is often illustrated by considering a stack of coins: If you put a number of coins on top of each other then the total volume of the coins is just the sum of the volumes of the coins. If you change the shape of the pile by sliding the coins horizontally then the volume of the pile will still be the sum of the volumes of the coins, i.e. it doesn't change.

3.4. Solids of revolution. In principle, formula (63) allows you to compute the volume of any solid, provided you can compute the areas $A(x)$ of all cross sections. One class of solids for which the areas of the cross sections are easy are the so-called “solids of revolution.”

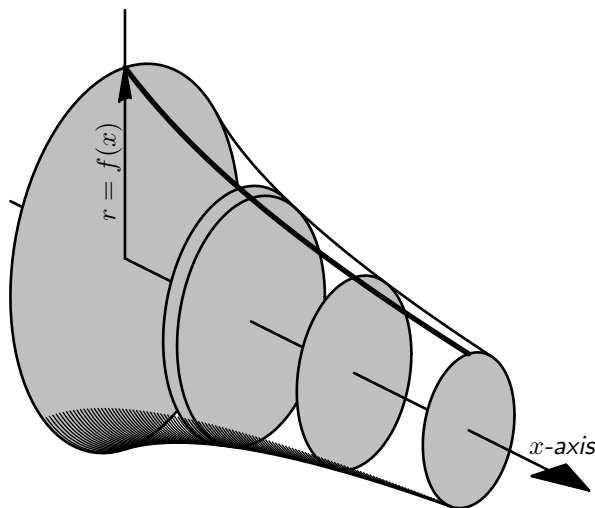


Figure 5. A solid of revolution consists of all points in three-dimensional space whose distance r to the x -axis satisfies $r \leq f(x)$.

A solid of revolution is created by rotating (revolving) the graph of a positive function around the x -axis. More precisely, let f be a function which is defined on an interval $[a, b]$ and which is always positive ($f(x) > 0$ for all x). If you now imagine the x -axis floating in three dimensional space, then the solid of revolution obtained by rotating the graph of f around the x -axis consists of all points in three-dimensional space with $a \leq x \leq b$, and whose distance to the x -axis is no more than $f(x)$.

Yet another way of describing the solid of revolution is to say that the solid is the union of all discs which meet the x -axis perpendicularly and whose radius is given by $r = f(x)$.

If we slice the solid with planes perpendicular to the x -axis, then (63) tells us the volume of the solid. Each slice is a disc of radius $r = f(x)$ so that its area is $A(x) = \pi r^2 = \pi f(x)^2$. We therefore find that

$$(64) \quad V = \pi \int_a^b f(x)^2 dx.$$

4. Examples of volumes of solids of revolution

4.1. Problem 1: Revolve \mathcal{R} around the y -axis . Consider the solid obtained by revolving the region

$$\mathcal{R} = \{(x, y) \mid 0 \leq x \leq 2, \quad (x-1)^2 \leq y \leq 1\}$$

around the y -axis.

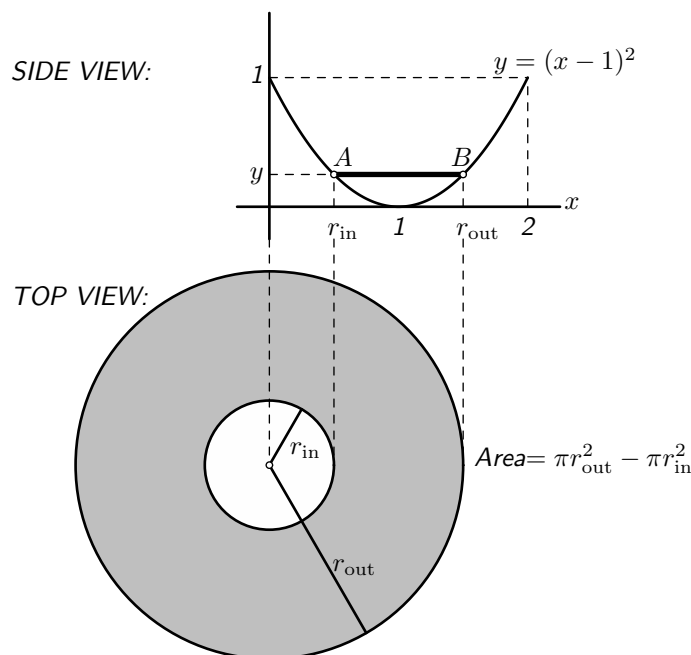
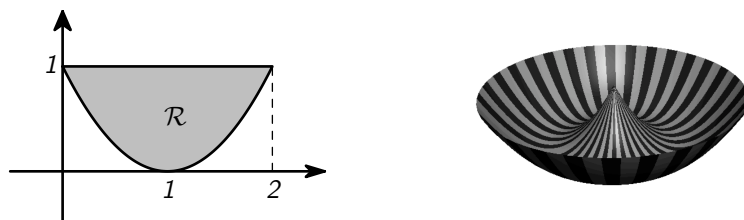


Figure 6. Computing the volume of the solid you get when you revolve the region \mathcal{R} around the y -axis. A horizontal cross section of the solid is a “washer” with inner radius r_{in} , and outer radius r_{out} .



Solution: The region we have to revolve around the y -axis consists of all points above the parabola $y = (x - 1)^2$ but below the line $y = 1$.

If we intersect the solid with a plane at height y then we get a ring shaped region, or “annulus”, i.e. a large disc with a smaller disc removed. You can see it in the figure below: if you cut the region \mathcal{R} horizontally at height y you get the line segment AB , and if you rotate this segment around the y -axis you get the grey ring region pictured below the graph. Call the radius of the outer circle r_{out} and the radius of the inner circle r_{in} . These radii are the two solutions of

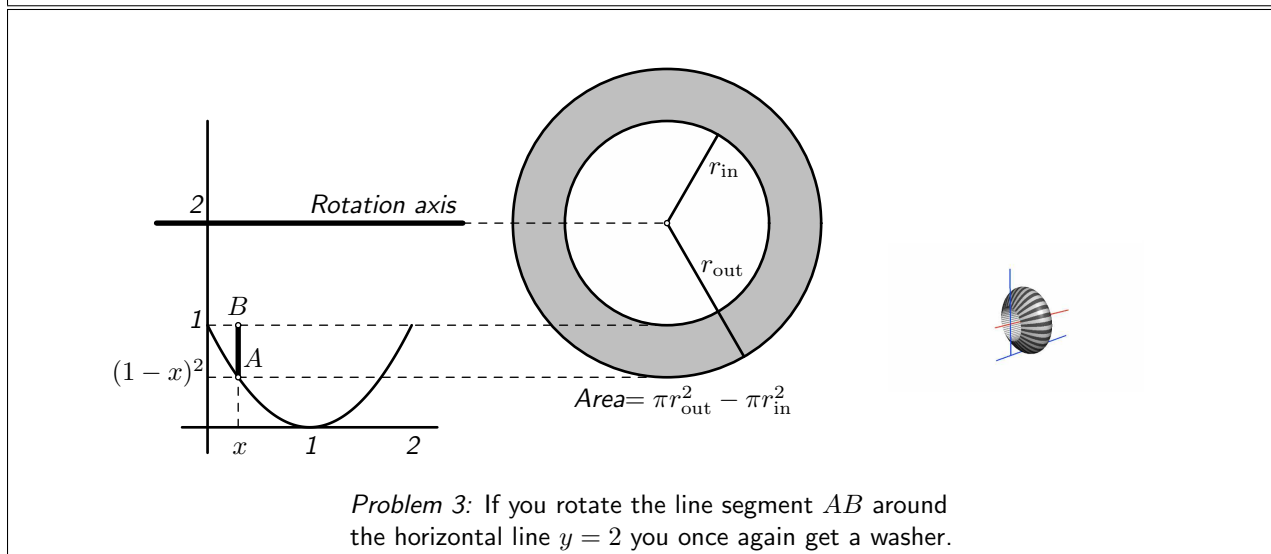
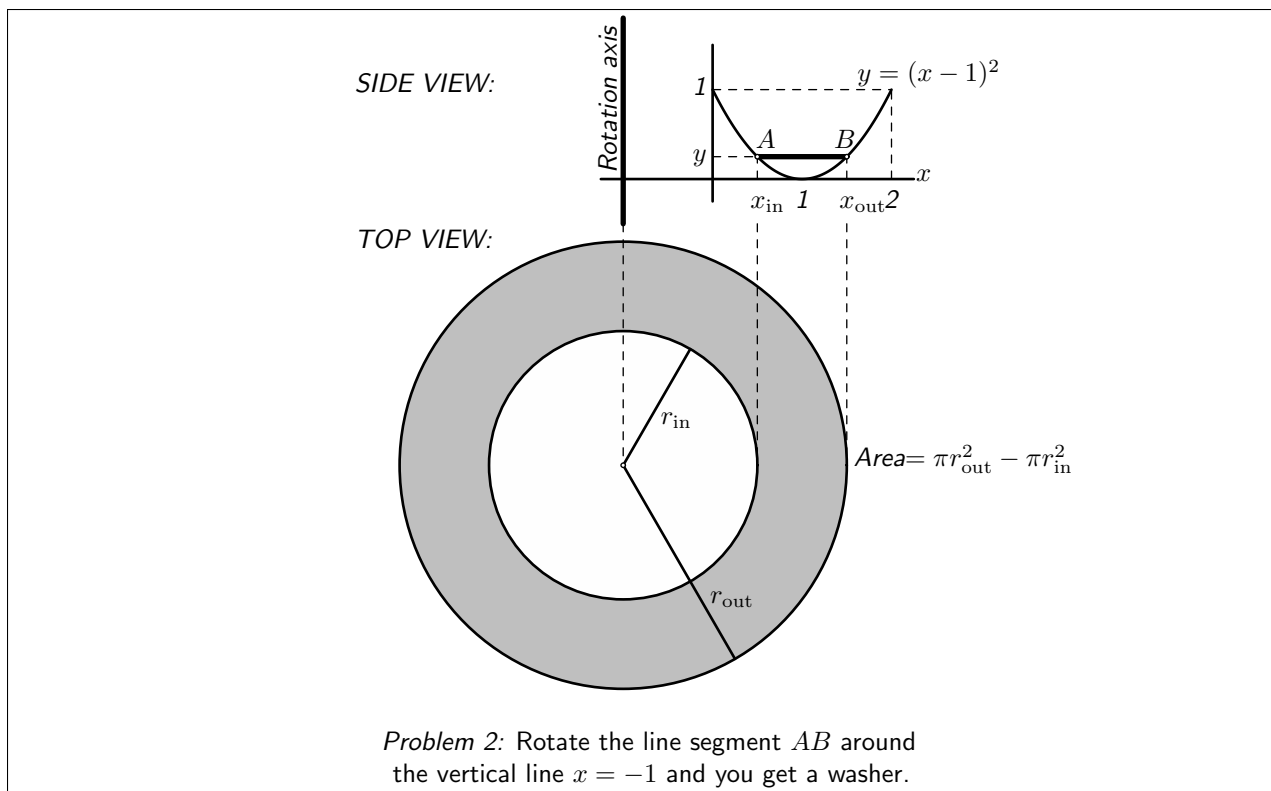
$$y = (1 - r)^2$$

so they are

$$r_{\text{in}} = 1 - \sqrt{y}, \quad r_{\text{out}} = 1 + \sqrt{y}.$$

The area of the cross section is therefore given by

$$A(y) = \pi r_{\text{out}}^2 - \pi r_{\text{in}}^2 = \pi(1 + \sqrt{y})^2 - \pi(1 - \sqrt{y})^2 = 4\pi\sqrt{y}.$$



The y -values which occur in the solid are $0 \leq y \leq 1$ and hence the volume of the solid is given by

$$V = \int_0^1 A(y) dy = 4\pi \int_0^1 \sqrt{y} dy = 4\pi \times \frac{2}{3} = \frac{8\pi}{3}.$$

4.2. Problem 2: Revolve \mathcal{R} around the line $x = -1$. Find the volume of the solid of revolution obtained by revolving the same region \mathcal{R} around the line $x = -1$.

Solution: The line $x = -1$ is vertical, so we slice the solid with horizontal planes. The height of each plane will be called y .

As before the slices are ring shaped regions but the inner and outer radii are now given by

$$r_{\text{in}} = 1 + x_{\text{in}} = 2 - \sqrt{y}, \quad r_{\text{out}} = 1 + x_{\text{out}} = 2 + \sqrt{y}.$$

The volume is therefore given by

$$V = \int_0^1 (\pi r_{\text{out}}^2 - \pi r_{\text{in}}^2) dy = \pi \int_0^1 8\sqrt{y} dy = \frac{16\pi}{3}.$$

4.3. Problem 3: Revolve \mathcal{R} around the line $y = 2$. Compute the volume of the solid you get when you revolve the same region \mathcal{R} around the line $y = 2$.

Solution: This time the line around which we rotate \mathcal{R} is horizontal, so we slice the solid with planes perpendicular to the x -axis.

A typical slice is obtained by revolving the line segment AB about the line $y = 2$. The result is again an annulus, and from the figure we see that the inner and outer radii of the annulus are

$$r_{\text{in}} = 1, \quad r_{\text{out}} = 2 - (1 - x)^2.$$

The area of the slice is therefore

$$A(x) = \pi \{2 - (1 - x)^2\}^2 - \pi 1^2 = \pi \{3 - 4(1 - x)^2 + (1 - x)^4\}.$$

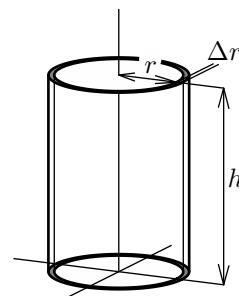
The x values which occur in the solid are $0 \leq x \leq 2$, and so its volume is

$$\begin{aligned} V &= \pi \int_0^2 \{3 - 4(1 - x)^2 + (1 - x)^4\} dx \\ &= \pi \left[3x + \frac{4}{3}(1 - x)^3 - \frac{1}{5}(1 - x)^5 \right]_0^2 \\ &= \frac{56}{15}\pi \end{aligned}$$

5. Volumes by cylindrical shells

Instead of slicing a solid with planes you can also try to decompose it into cylindrical shells. The volume of a cylinder of height h and radius r is $\pi r^2 h$ (height times area base). Therefore the volume of a cylindrical shell of height h , (inner) radius r and thickness Δr is

$$\begin{aligned} \pi h(r + \Delta r)^2 - \pi h r^2 &= \pi h(2r + \Delta r)\Delta r \\ &\approx 2\pi h r \Delta r. \end{aligned}$$



Now consider the solid you get by revolving the region

$$\mathcal{R} = \{(x, y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\}$$

around the y -axis. By partitioning the interval $a \leq x \leq b$ into many small intervals we can decompose the solid into many thin shells. The volume of each shell will approximately be given by $2\pi x f(x) \Delta x$. Adding the volumes of the shells, and taking the limit over finer and finer partitions we arrive at the following formula for the volume of the solid of revolution:

$$(65) \quad V = 2\pi \int_a^b x f(x) dx.$$

If the region \mathcal{R} is not the region under the graph, but rather the region between the graphs of two functions $f(x) \leq g(x)$, then we get

$$V = 2\pi \int_a^b x \{g(x) - f(x)\} dx.$$

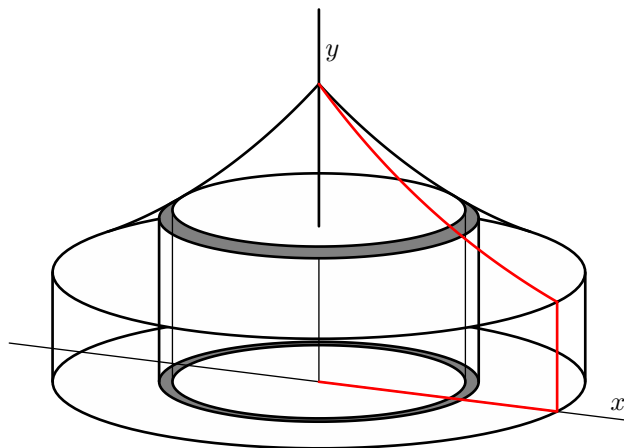


Figure 7. Computing the volume of a circus tent using cylindrical shells. This particular tent is obtained by rotating the graph of $y = e^{-x}$, $0 \leq x \leq 1$ around the y -axis.

5.1. Example – The solid obtained by rotating \mathcal{R} about the y -axis, again. The region \mathcal{R} from §4.1 can also be described as

$$\mathcal{R} = \{(x, y) \mid 0 \leq x \leq 2, f(x) \leq y \leq g(x)\},$$

where

$$f(x) = (x-1)^2 \text{ and } g(x) = 1.$$

The volume of the solid which we already computed in §4.1 is thus given by

$$\begin{aligned} V &= 2\pi \int_0^1 x \{1 - (x-1)^2\} dx \\ &= 2\pi \int_0^2 \{-x^3 + 2x^2\} dx \\ &= 2\pi \left[-\frac{1}{4}x^4 + \frac{2}{3}x^3\right]_0^2 \\ &= 8\pi/3, \end{aligned}$$

which coincides with the answer we found in §4.1.

6. Exercises

475. Group Problem.

What do the dots in “lim...” in equation (62) stand for? (i.e. what approaches what in this limit?)

Draw and describe the solids whose volume you are asked to compute in the following problems:

476. Find the volume enclosed by the paraboloid obtained by rotating the graph of $f(x) = R\sqrt{x/H}$ ($0 \leq x \leq H$) around the x -axis. Here R and H are positive constants. Draw the solid whose volume you are asked to compute, and indicate what R and H are in your drawing.

Find the volume of the solids you get by rotating each of the following graphs around the x -axis:

477. $f(x) = x$, $0 \leq x \leq 2$

478. $f(x) = \sqrt{2-x}$, $0 \leq x \leq 2$

479. $f(x) = (1+x^2)^{-1/2}$, $|x| \leq 1$

480. $f(x) = \sin x$, $0 \leq x \leq \pi$

481. $f(x) = 1-x^2$, $|x| \leq 1$

482. $f(x) = \cos x$, $0 \leq x \leq \pi$ (!!)

483. $f(x) = 1/\cos x$, $0 \leq x \leq \pi/4$

484. Find the volume that results by rotating the semicircle $y = \sqrt{R^2 - x^2}$ about the x -axis.

485. Let \mathcal{T} be triangle $1 \leq x \leq 2$, $0 \leq y \leq 3x - 3$.

(a) Find the volume of the solid obtained by rotating the triangle \mathcal{T} around the x -axis.

(b) Find the volume that results by rotating the triangle \mathcal{T} around the y axis.

(c) Find the volume that results by rotating the triangle \mathcal{T} around the line $x = -1$.

(d) Find the volume that results by rotating the triangle \mathcal{T} around the line $y = -1$.

486. Group Problem.

(a) A spherical bowl of radius a contains water to a depth $h < 2a$. Find the volume of the water in the bowl. (Which solid of revolution is implied in this problem?)

(b) Water runs into a spherical bowl of radius 5 ft at the rate of $0.2 \text{ ft}^3/\text{sec}$. How fast is the water level rising when the water is 4 ft deep?

7. Distance from velocity, velocity from acceleration

7.1. Motion along a line. If an object is moving on a straight line, and if its position at time t is $x(t)$, then we had defined the velocity to be $v(t) = x'(t)$. Therefore the position is an antiderivative of the velocity, and the fundamental theorem of calculus says that

$$(66) \quad \int_{t_a}^{t_b} v(t) dt = x(t_b) - x(t_a),$$

or

$$x(t_b) = x(t_a) + \int_{t_a}^{t_b} v(t) dt.$$

In words, the integral of the velocity gives you the distance travelled of the object (during the interval of integration).

Equation (66) can also be obtained using Riemann sums. Namely, to see how far the object moved between times t_a and t_b we choose a partition $t_a = t_0 < t_1 < \cdots < t_N = t_b$. Let Δs_k be the distance travelled during the time interval (t_{k-1}, t_k) . The length of this time interval is $\Delta t_k = t_k - t_{k-1}$. During this time interval the velocity $v(t)$ need not be constant, but if the time interval is short enough then we can estimate the velocity by $v(c_k)$ where c_k is some number between t_{k-1} and t_k . We then have

$$\Delta s_k = v(c_k) \Delta t_k$$

and hence the total distance travelled is the sum of the travel distances for all time intervals $t_{k-1} < t < t_k$, i.e.

$$\text{Distance travelled} \approx \Delta s_1 + \cdots + \Delta s_N = v(c_1) \Delta t_1 + \cdots + v(c_N) \Delta t_N.$$

The right hand side is again a Riemann sum for the integral in (66). As one makes the partition finer and finer you therefore get

$$\text{Distance travelled} = \int_{t_a}^{t_b} v(t) dt.$$

The return of the dummy. Often you want to write a formula for $x(t) = \cdots$ rather than $x(t_b) = \cdots$ as we did in (66), i.e. you want to say what the position is at time t , instead of at time t_a . For instance, you might want to express the fact that the position $x(t)$ is equal to the initial position $x(0)$ plus the integral of the velocity from 0 to t . To do this you cannot write

$$x(t) = x(0) + \int_0^t v(t) dt \quad \Leftarrow \quad \text{BAD FORMULA}$$

because the variable t gets used in two incompatible ways: the t in $x(t)$ on the left, and in the upper bound on the integral (\int^t) are the same, but they are not the same as the two t 's in $v(t)dt$. The latter is a dummy variable (see §8 and §3.3). To fix this formula we should choose a different letter or symbol for the integration variable. A common choice in this situation is to decorate the integration variable with a prime (t'), a tilde (\bar{t}) or a bar (\bar{t}). So you can write

$$x(t) = x(0) + \int_0^t v(\bar{t}) d\bar{t}$$

7.2. Velocity from acceleration. The acceleration of the object is by definition the rate of change of its velocity,

$$a(t) = v'(t),$$

so you have

$$v(t) = v(0) + \int_0^t a(\bar{t}) d\bar{t}.$$

Conclusion: *If you know the acceleration $a(t)$ at all times t , and also the velocity $v(0)$ at time $t = 0$, then you can compute the velocity $v(t)$ at all times by integrating.*

7.3. Free fall in a constant gravitational field. If you drop an object then it will fall, and as it falls its velocity increases. The object's motion is described by the fact that *its acceleration is constant*. This constant is called g and is about $9.8\text{m/sec}^2 \approx 32\text{ft/sec}^2$. If we designate the upward direction as positive then $v(t)$ is the upward velocity of the object, and this velocity is actually decreasing. Therefore the constant acceleration is negative: it is $-g$.

If you write $h(t)$ for the height of the object at time t then its velocity is $v(t) = h'(t)$, and its acceleration is $h''(t)$. Since the acceleration is constant you have the following formula for the velocity at time t :

$$v(t) = v(0) + \int_0^t (-g) d\bar{t} = v(0) - gt.$$

Here $v(0)$ is the velocity at time $t = 0$ (the "initial velocity").

To get the height of the object at any time t you must integrate the velocity:

$$\begin{aligned} h(t) &= h(0) + \int_0^t v(\bar{t}) d\bar{t} && \text{(Note the use of the dummy } \bar{t} \text{)} \\ &= h(0) + \int_0^t \{v(0) - g\bar{t}\} d\bar{t} && \text{(use } v(\bar{t}) = v(0) - g\bar{t} \text{)} \\ &= h(0) + [v(0)\bar{t} - \frac{1}{2}g\bar{t}^2]_{\bar{t}=0}^{\bar{t}=t} \\ &= h(0) + v(0)t - \frac{1}{2}gt^2. \end{aligned}$$

For instance, if you launch the object upwards with velocity 5ft/sec from a height of 10ft, then you have

$$h(0) = 10\text{ft}, \quad v(0) = +5\text{ft/sec},$$

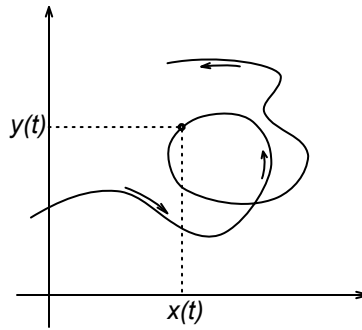
and thus

$$h(t) = 10 + 5t - 32t^2/2 = 10 + 5t - 16t^2.$$

The object reaches its maximum height when $h(t)$ has a maximum, which is when $h'(t) = 0$. To find that height you compute $h'(t) = 5 - 32t$ and conclude that $h(t)$ is maximal at $t = \frac{5}{32}\text{sec}$. The maximal height is then

$$h_{\max} = h(\frac{5}{32}) = 10 + \frac{25}{32} - \frac{25}{64} = 10\frac{25}{64}\text{ft}.$$

7.4. Motion in the plane – parametric curves. To describe the motion of an object in the plane you could keep track of its x and y coordinates at all times t . This would give you *two* functions of t , namely, $x(t)$ and $y(t)$, both of which are defined on the same interval $t_0 \leq t \leq t_1$ which describes the duration of the motion you are describing. In this context a pair of functions $(x(t), y(t))$ is called **a parametric curve**.



As an example, consider the motion described by

$$x(t) = \cos t, \quad y(t) = \sin t \quad (0 \leq t \leq 2\pi).$$

In this motion the point $(x(t), y(t))$ lies on the unit circle since

$$x(t)^2 + y(t)^2 = \cos^2 t + \sin^2 t = 1.$$

As t increases from 0 to 2π the point $(x(t), y(t))$ goes around the unit circle exactly once, in the counter-clockwise direction.

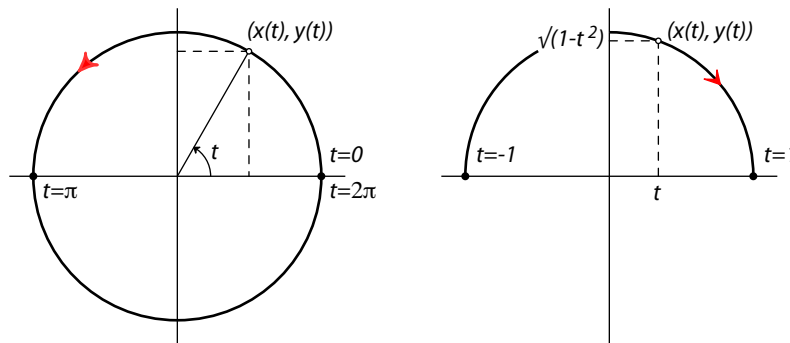


Figure 8. Two motions in the plane. On the left $x(t) = \cos t$, $y(t) = \sin t$ with $0 \leq t \leq 2\pi$, and on the right $x(t) = t$, $y(t) = \sqrt{1-t^2}$ with $-1 \leq t \leq 1$.

In another example one could consider

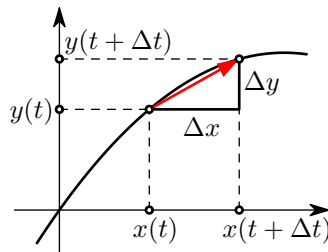
$$x(t) = t, \quad y(t) = \sqrt{1-t^2}, \quad (-1 \leq t \leq 1).$$

Here at all times the x and y coordinates satisfy

$$x(t)^2 + y(t)^2 = 1$$

again so that the point $(x(t), y(t)) = (t, \sqrt{1-t^2})$ again lies on the unit circle. Unlike the previous example we now always have $y(t) \geq 0$ (since $y(t)$ is the square root of something), and unlike the previous example the motion is only defined for $-1 \leq t \leq 1$. As t increases from -1 to $+1$, $x(t) = t$ does the same, and hence the point $(x(t), y(t))$ moves along the *upper half of the unit circle* from the leftmost point to the rightmost point.

7.5. The velocity of an object moving in the plane. We have seen that the velocity of an object which is moving along a line is the derivative of its position. If the object is allowed to move in the plane, so that its motion is described by a parametric curve $(x(t), y(t))$, then we can differentiate both $x(t)$ and $y(t)$, which gives us $x'(t)$ and $y'(t)$, and which leaves us with the following question: *At what speed is a particle moving if it is undergoing the motion $(x(t), y(t))$ ($t_a \leq t \leq t_b$) ?*



To answer this question we consider a short time interval $(t, t + \Delta t)$. During this time interval the particle moves from $(x(t), y(t))$ to $(x(t + \Delta t), y(t + \Delta t))$. Hence it has traveled a distance

$$\Delta s = \sqrt{(\Delta x)^2 + (\Delta y)^2}$$

where

$$\Delta x = x(t + \Delta t) - x(t), \text{ and } \Delta y = y(t + \Delta t) - y(t).$$

Dividing by Δt you get

$$\frac{\Delta s}{\Delta t} = \sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2},$$

for the average velocity over the time interval $[t, t + \Delta t]$. Letting $\Delta t \rightarrow 0$ you find the velocity at time t to be

$$(67) \quad v = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}.$$

7.6. Example – the two motions on the circle from §7.4. If a point moves along a circle according to $x(t) = \cos t$, $y(t) = \sin t$ (figure 8 on the left) then

$$\frac{dx}{dt} = -\sin t, \quad \frac{dy}{dt} = +\cos t$$

so

$$v(t) = \sqrt{(-\cos t)^2 + (\sin t)^2} = 1.$$

The velocity of this motion is therefore always the same; the point $(\cos t, \sin t)$ moves along the unit circle with constant velocity.

In the second example in §7.4 we had $x(t) = t$, $y(t) = \sqrt{1 - t^2}$, so

$$\frac{dx}{dt} = 1, \quad \frac{dy}{dt} = \frac{-t}{\sqrt{1 - t^2}}$$

whence

$$v(t) = \sqrt{1^2 + \frac{t^2}{1 - t^2}} = \frac{1}{\sqrt{1 - t^2}}.$$

Therefore the point $(t, \sqrt{1 - t^2})$ moves along the upper half of the unit circle from the left to the right, and its velocity changes according to $v = 1/\sqrt{1 - t^2}$.

8. The length of a curve

8.1. Length of a parametric curve. Let $(x(t), y(t))$ be some parametric curve defined for $t_a \leq t \leq t_b$. To find the length of this curve you can reason as follows: The length of the curve should be the distance travelled by the point $(x(t), y(t))$ as t increases from t_a to t_b . At each moment in time the velocity $v(t)$ of the point is given by (67), and therefore the distance traveled should be

$$(68) \quad s = \int_{t_a}^{t_b} v(t) dt = \int_{t_a}^{t_b} \sqrt{x'(t)^2 + y'(t)^2} dt.$$

Alternatively, you could try to compute the distance travelled by means of Riemann sums. Choose a partition

$$t_a = t_0 < t_1 < \cdots < t_N = t_b$$

of the interval $[t_a, t_b]$. You then get a sequence of points $P_0(x(t_0), y(t_0))$, $P_1(x(t_1), y(t_1))$, \dots , $P_N(x(t_N), y(t_N))$, and after “connecting the dots” you get a polygon. You could approximate the length of the curve by computing the length of this polygon. The distance between two consecutive points P_{k-1} and P_k is

$$\begin{aligned}\Delta s_k &= \sqrt{(\Delta x_k)^2 + (\Delta y_k)^2} \\ &= \sqrt{\left(\frac{\Delta x_k}{\Delta t_k}\right)^2 + \left(\frac{\Delta y_k}{\Delta t_k}\right)^2} \Delta t_k \\ &\approx \sqrt{x'(c_k)^2 + y'(c_k)^2} \Delta t_k\end{aligned}$$

where we have approximated the difference quotients

$$\frac{\Delta x_k}{\Delta t_k} \text{ and } \frac{\Delta y_k}{\Delta t_k}$$

by the derivatives $x'(c_k)$ and $y'(c_k)$ for some c_k in the interval $[t_{k-1}, t_k]$.

The total length of the polygon is then

$$\sqrt{x'(c_1)^2 + y'(c_1)^2} \Delta t_1 + \cdots + \sqrt{x'(c_N)^2 + y'(c_N)^2} \Delta t_N$$

This is a Riemann sum for the integral $\int_{t_a}^{t_b} \sqrt{x'(t)^2 + y'(t)^2} dt$, and hence we find (once more) that the length of the curve is

$$s = \int_{t_a}^{t_b} \sqrt{x'(t)^2 + y'(t)^2} dt.$$

8.2. The length of the graph of a function. The graph of a function ($y = f(x)$ with $a \leq x \leq b$) is also a curve in the plane, and you can ask what its length is. We will now find this length by *representing the graph as a parametric curve* and applying the formula (68) from the previous section.

The standard method of representing the graph of a function $y = f(x)$ by a parametric curve is to choose

$$x(t) = t, \text{ and } y(t) = f(t), \text{ for } a \leq t \leq b.$$

This parametric curve traces the graph of $y = f(x)$ from left to right as t increases from a to b .

Since $x'(t) = 1$ and $y'(t) = f'(t)$ we find that the length of the graph is

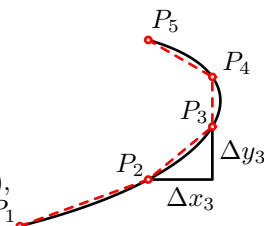
$$L = \int_a^b \sqrt{1 + f'(t)^2} dt.$$

The variable t in this integral is a dummy variable and we can replace it with any other variable we like, for instance, x :

$$(69) \quad L = \int_a^b \sqrt{1 + f'(x)^2} dx.$$

In Leibniz' notation we have $y = f(x)$ and $f'(x) = dy/dx$ so that Leibniz would have written

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx.$$



9. Examples of length computations

9.1. Length of a circle. In §8 we parametrized the unit circle by

$$x(t) = \cos t, \quad y(t) = \sin t, \quad (0 \leq t \leq 2\pi)$$

and computed $\sqrt{x'(t)^2 + y'(t)^2} = 1$. Therefore our formula tells us that the length of the unit circle is

$$L = \int_0^{2\pi} \sqrt{x'(t)^2 + y'(t)^2} dt = \int_0^{2\pi} 1 dt = 2\pi.$$

This cannot be a PROOF that the unit circle has length 2π since we have already used that fact to define angles in radians, to define the trig functions Sine and Cosine, and to find their derivatives. But our computation shows that the length formula (69) is at least consistent with what we already knew.

9.2. Length of a parabola. Consider our old friend, the parabola $y = x^2$, $0 \leq x \leq 1$. While the area under its graph was easy to compute ($\frac{1}{3}$), its length turns out to be much more complicated.

Our length formula (69) says that the length of the parabola is given by

$$L = \int_0^1 \sqrt{1 + \left(\frac{dx^2}{dx}\right)^2} dx = \int_0^1 \sqrt{1 + 4x^2} dx.$$

To find this integral you would have to use one of the following (not at all obvious) substitutions¹

$$x = \frac{1}{4}\left(z - \frac{1}{z}\right) \quad (\text{then } 1 + 4x^2 = \frac{1}{4}(z + 1/z)^2 \text{ so you can simplify the } \sqrt{\cdot})$$

or (if you like hyperbolic functions)

$$x = \frac{1}{2} \sinh w \quad (\text{in which case } \sqrt{1 + 4x^2} = \cosh w.)$$

9.3. Length of the graph of the Sine function. To compute the length of the curve given by $y = \sin x$, $0 \leq x \leq \pi$ you would have to compute this integral:

$$(70) \quad L = \int_0^\pi \sqrt{1 + \left(\frac{d \sin x}{dx}\right)^2} dx = \int_0^\pi \sqrt{1 + \cos^2 x} dx.$$

Unfortunately this is not an integral which can be computed in terms of the functions we know in this course (it's an "elliptic integral of the second kind.") This happens very often with the integrals that you run into when you try to compute the length of a curve. In spite of the fact that we get stuck when we try to compute the integral in (70), the formula is not useless. For example, since $-1 \leq \cos x \leq 1$ we know that

$$1 \leq \sqrt{1 + \cos^2 x} \leq \sqrt{1 + 1} = \sqrt{2},$$

and therefore the length of the Sine graph is bounded by

$$\int_0^\pi 1 dx \leq \int_0^\pi \sqrt{1 + \cos^2 x} dx \leq \int_0^\pi \sqrt{2} dx,$$

i.e.

$$\pi \leq L \leq \pi\sqrt{2}.$$

10. Exercises

487. Find the length of the piece of the graph of $y = \sqrt{1 - x^2}$ where $0 \leq x \leq \frac{1}{2}$.

The graph is a circle, so there are two ways of computing this length. One uses geometry (length of a circular arc = radius \times angle), the other uses an integral.

Use both methods and check that you get the same answer.

488. Compute the length of the part of the *evolute* of the circle, given by

$$x(t) = \cos t - t \sin t, \quad y(t) = \sin t + t \cos t$$

where $0 < t < \pi$.

489. Group Problem.

¹Many calculus textbooks will tell you to substitute $x = \tan \theta$, but the resulting integral is still not easy.

Show that the *Archimedean spiral*, given by

$$x(\theta) = \theta \cos \theta, \quad y(\theta) = \theta \sin \theta, \quad 0 \leq \theta \leq \pi$$

has the same length as the parabola given by

$$y = \frac{1}{2}x^2, \quad 0 \leq x \leq \pi.$$

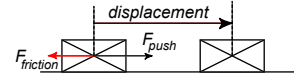
Hint: you can set up integrals for both lengths. If you get the same integral in both cases, then you know the two curves have the same length (even if you don't try to compute the integrals).

11. Work done by a force

11.1. Work as an integral. In Newtonian mechanics a force which acts on an object in motion performs a certain amount of work, i.e. it spends a certain amount of energy. If the force which acts is constant, then the work done by this force is

$$\text{Work} = \text{Force} \times \text{Displacement}.$$

For example if you are pushing a box forward then there will be two forces acting on the box: the force you apply, and the friction force of the floor on the box. The amount of work you do is the product of the force you exert and the length of the displacement. Both displacement and the force you apply are pointed towards the right, so both are positive, and the work you do (energy you provide to the box) is positive.



The amount of work done by the friction is similarly the product of the friction force and the displacement. Here the displacement is still to the right, but the friction force points to the left, so it is negative. The work done by the friction force is therefore negative. Friction extracts energy from the box.

Suppose now that the force $F(t)$ on the box is not constant, and that its motion is described by saying that its position at time t is $x(t)$. The basic formula work = force \times displacement does not apply directly since it assumes that the force is constant. To compute the work done by the varying force $F(t)$ we choose a partition of the time interval $t_a \leq t \leq t_b$ into

$$t_a = t_0 < t_1 < \cdots < t_{N-1} < t_N = t_b$$

In each short time interval $t_{k-1} \leq t \leq t_k$ we assume the force is (almost) constant and we approximate it by $F(c_k)$ for some $t_{k-1} \leq c_k \leq t_k$. If we also assume that the velocity $v(t) = x'(t)$ is approximately constant between times t_{k-1} and t_k then the displacement during this time interval will be

$$x(t_k) - x(t_{k-1}) \approx v(c_k)\Delta t_k,$$

where $\Delta t_k = t_k - t_{k-1}$. Therefore the work done by the force F during the time interval $t_{k-1} \leq t \leq t_k$ is

$$\Delta W_k = F(c_k)v(c_k)\Delta t_k.$$

Adding the work done during each time interval we get the total work done by the force between time t_a and t_b :

$$W = F(c_1)v(c_1)\Delta t_1 + \cdots + F(c_N)v(c_N)\Delta t_N.$$

Again we have a Riemann sum for an integral. If we take the limit over finer and finer partitions we therefore find that the work done by the force $F(t)$ on an object whose motion is described by $x(t)$ is

$$(71) \quad W = \int_{t_a}^{t_b} F(t)v(t)dt,$$

in which $v(t) = x'(t)$ is the velocity of the object.

11.2. Kinetic energy. Newton's famous law relating the force exerted on an object and its motion says $F = ma$, where a is the acceleration of the object, m is its mass, and F is the combination of all forces acting on the object. If the position of the object at time t is $x(t)$, then its velocity and acceleration are $v(t) = x'(t)$ and $a(t) = v'(t) = x''(t)$, and thus the total force acting on the object is

$$F(t) = ma(t) = m \frac{dv}{dt}.$$

The work done by the total force is therefore

$$(72) \quad W = \int_{t_a}^{t_b} F(t)v(t)dt = \int_{t_a}^{t_b} m \frac{dv(t)}{dt} v(t) dt.$$

Even though we have not assumed anything about the motion, so we don't know anything about the velocity $v(t)$, we can still do this integral. The key is to notice that, by the chain rule,

$$m \frac{dv(t)}{dt} v(t) = \frac{d}{dt} \frac{1}{2} m v(t)^2.$$

(Remember that m is a constant.) This says that the quantity

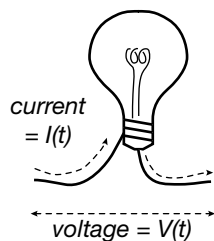
$$K(t) = \frac{1}{2} m v(t)^2$$

is the antiderivative we need to do the integral (72). We get

$$W = \int_{t_a}^{t_b} m \frac{dv(t)}{dt} v(t) dt = \int_{t_a}^{t_b} K'(t) dt = K(t_b) - K(t_a).$$

In Newtonian mechanics the quantity $K(t)$ is called **the kinetic energy** of the object, and our computation shows that *the amount by which the kinetic energy of an object increases is equal to the amount of work done on the object.*

12. Work done by an electric current



If at time t an electric current $I(t)$ (measured in Ampère) flows through an electric circuit, and if the voltage across this circuit is $V(t)$ (measured in Volts) then the energy supplied to the circuit per second is $I(t)V(t)$. Therefore the total energy supplied during a time interval $t_0 \leq t \leq t_1$ is the integral

$$\text{Energy supplied} = \int_{t_0}^{t_1} I(t)V(t)dt.$$

(measured in Joule; the energy consumption of a circuit is defined to be how much energy it consumes per time unit, and the power consumption of a circuit which consumes 1 Joule per second is said to be one Watt.)

If a certain voltage is applied to a simple circuit (like a light bulb) then the current flowing through that circuit is determined by the resistance R of that circuit by *Ohm's law*² which says

$$I = \frac{V}{R}.$$

²http://en.wikipedia.org/wiki/Ohm's_law

12.1. Example. If the resistance of a light bulb is $R = 200\Omega$, and if the voltage applied to it is

$$V(t) = 150 \sin 2\pi ft$$

where $f = 50\text{sec}^{-1}$ is the frequency, then how much energy does the current supply to the light bulb in one second?

To compute this we first find the current using Ohm's law,

$$I(t) = \frac{V(t)}{R} = \frac{150}{200} \sin 2\pi ft = 0.75 \sin 2\pi ft. \quad (\text{Amp})$$

The energy supplied in one second is then

$$\begin{aligned} E &= \int_0^{1 \text{ sec}} I(t)V(t)dt \\ &= \int_0^1 (150 \sin 2\pi ft) \times (0.75 \sin 2\pi ft) dt \\ &= 112.5 \int_0^1 \sin^2(2\pi ft) dt \end{aligned}$$

You can do this last integral by using the double angle formula for the cosine, to rewrite

$$\sin^2(2\pi ft) = \frac{1}{2}\{1 - \cos 4\pi ft\} = \frac{1}{2} - \frac{1}{2} \cos 4\pi ft.$$

Keep in mind that $f = 50$, and you find that the integral is

$$\int_0^1 \sin^2(2\pi ft) dt = \left[\frac{t}{2} - \frac{1}{4\pi f} \sin 4\pi ft \right]_0^1 = \frac{1}{2},$$

and hence the energy supplied to the light bulb during one second is

$$E = 112.5 \times \frac{1}{2} = 56.25(\text{Joule}).$$

Chapter seven:

Line and surface integrals functions

1.6 Surfaces

In the previous section we discussed planes in Euclidean space. A plane is an example of a *surface*, which we will define informally⁸ as the solution set of the equation $F(x, y, z) = 0$ in \mathbb{R}^3 , for some real-valued function F . For example, a plane given by $ax + by + cz + d = 0$ is the solution set of $F(x, y, z) = 0$ for the function $F(x, y, z) = ax + by + cz + d$. Surfaces are 2-dimensional. The plane is the simplest surface, since it is “flat”. In this section we will look at some surfaces that are more complex, the most important of which are the sphere and the cylinder.

Definition 1.9. A sphere S is the set of all points (x, y, z) in \mathbb{R}^3 which are a fixed distance r (called the **radius**) from a fixed point $P_0 = (x_0, y_0, z_0)$ (called the **center** of the sphere):

$$S = \{(x, y, z) : (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2\} \quad (1.29)$$

Using vector notation, this can be written in the equivalent form:

$$S = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}_0\| = r\} \quad (1.30)$$

where $\mathbf{x} = (x, y, z)$ and $\mathbf{x}_0 = (x_0, y_0, z_0)$ are vectors.

Figure 1.6.1 illustrates the vectorial approach to spheres.

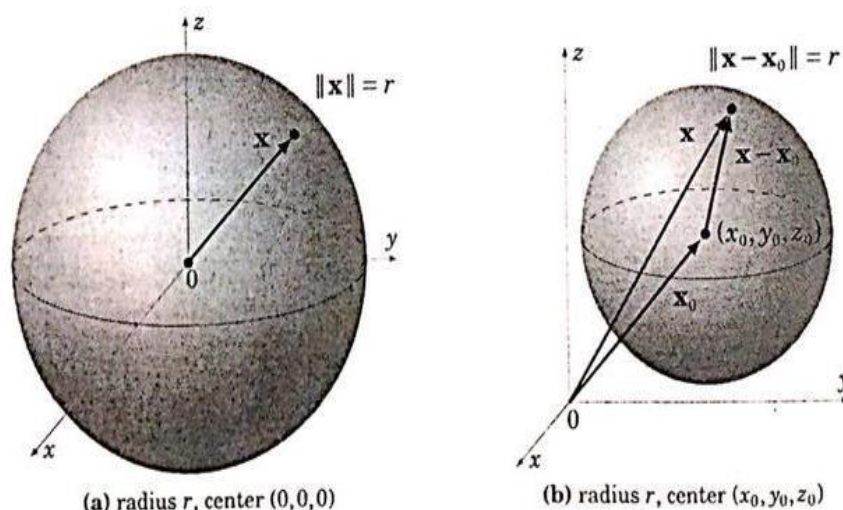


Figure 1.6.1 Spheres in \mathbb{R}^3

Note in Figure 1.6.1(a) that the intersection of the sphere with the xy -plane is a circle of radius r (i.e. a *great circle*, given by $x^2 + y^2 = r^2$ as a subset of \mathbb{R}^2). Similarly for the intersections with the xz -plane and the yz -plane. In general, a plane intersects a sphere either at a single point or in a circle.

⁸See O'NEILL for a deeper and more rigorous discussion of surfaces.

Example 1.27. Find the intersection of the sphere $x^2 + y^2 + z^2 = 169$ with the plane $z = 12$.

Solution: The sphere is centered at the origin and has radius $13 = \sqrt{169}$, so it does intersect the plane $z = 12$. Putting $z = 12$ into the equation of the sphere gives

$$\begin{aligned}x^2 + y^2 + 12^2 &= 169 \\x^2 + y^2 &= 169 - 144 = 25 = 5^2\end{aligned}$$

which is a circle of radius 5 centered at $(0, 0, 12)$, parallel to the xy -plane (see Figure 1.6.2).

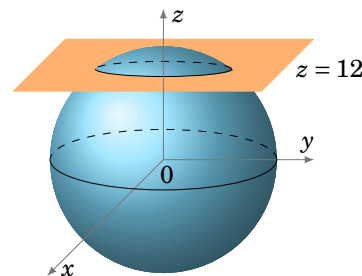


Figure 1.6.2

If the equation in formula (1.29) is multiplied out, we get an equation of the form:

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0 \quad (1.31)$$

for some constants a , b , c and d . Conversely, an equation of this form *may* describe a sphere, which can be determined by completing the square for the x , y and z variables.

Example 1.28. Is $2x^2 + 2y^2 + 2z^2 - 8x + 4y - 16z + 10 = 0$ the equation of a sphere?

Solution: Dividing both sides of the equation by 2 gives

$$\begin{aligned}x^2 + y^2 + z^2 - 4x + 2y - 8z + 5 &= 0 \\(x^2 - 4x + 4) + (y^2 + 2y + 1) + (z^2 - 8z + 16) + 5 - 4 - 1 - 16 &= 0 \\(x - 2)^2 + (y + 1)^2 + (z - 4)^2 &= 16\end{aligned}$$

which is a sphere of radius 4 centered at $(2, -1, 4)$.

Example 1.29. Find the point(s) of intersection (if any) of the sphere from Example 1.28 and the line $x = 3 + t$, $y = 1 + 2t$, $z = 3 - t$.

Solution: Put the equations of the line into the equation of the sphere, which was $(x - 2)^2 + (y + 1)^2 + (z - 4)^2 = 16$, and solve for t :

$$\begin{aligned}(3 + t - 2)^2 + (1 + 2t + 1)^2 + (3 - t - 4)^2 &= 16 \\(t + 1)^2 + (2t + 2)^2 + (-t - 1)^2 &= 16 \\6t^2 + 12t - 10 &= 0\end{aligned}$$

The quadratic formula gives the solutions $t = -1 \pm \frac{4}{\sqrt{6}}$. Putting those two values into the equations of the line gives the following two points of intersection:

$$\left(2 + \frac{4}{\sqrt{6}}, -1 + \frac{8}{\sqrt{6}}, 4 - \frac{4}{\sqrt{6}}\right) \quad \text{and} \quad \left(2 - \frac{4}{\sqrt{6}}, -1 - \frac{8}{\sqrt{6}}, 4 + \frac{4}{\sqrt{6}}\right)$$

If two spheres intersect, they do so either at a single point or in a circle.

Example 1.30. Find the intersection (if any) of the spheres $x^2 + y^2 + z^2 = 25$ and $x^2 + y^2 + (z - 2)^2 = 16$.

Solution: For any point (x, y, z) on both spheres, we see that

$$\begin{aligned} x^2 + y^2 + z^2 = 25 &\Rightarrow x^2 + y^2 = 25 - z^2, \text{ and} \\ x^2 + y^2 + (z - 2)^2 = 16 &\Rightarrow x^2 + y^2 = 16 - (z - 2)^2, \text{ so} \\ 16 - (z - 2)^2 = 25 - z^2 &\Rightarrow 4z - 4 = 9 \Rightarrow z = 13/4 \\ &\Rightarrow x^2 + y^2 = 25 - (13/4)^2 = 231/16 \end{aligned}$$

\therefore The intersection is the circle $x^2 + y^2 = \frac{231}{16}$ of radius $\frac{\sqrt{231}}{4} \approx 3.8$ centered at $(0, 0, \frac{13}{4})$.

The cylinders that we will consider are *right circular cylinders*. These are cylinders obtained by moving a line L along a circle C in \mathbb{R}^3 in a way so that L is always perpendicular to the plane containing C . We will only consider the cases where the plane containing C is parallel to one of the three coordinate planes (see Figure 1.6.3).

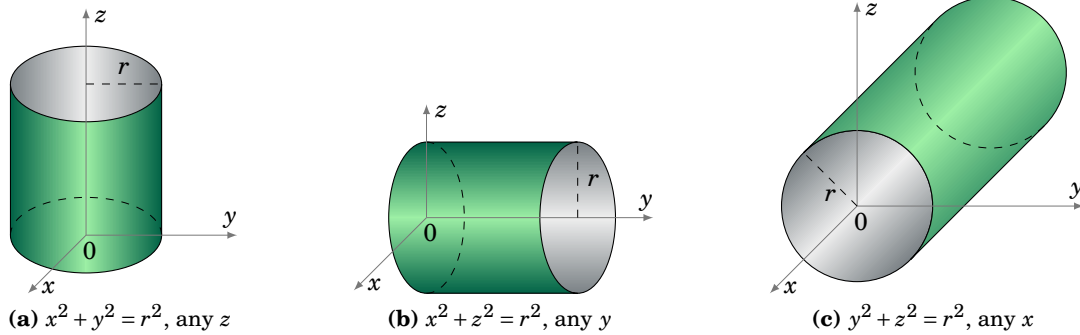


Figure 1.6.3 Cylinders in \mathbb{R}^3

For example, the equation of a cylinder whose base circle C lies in the xy -plane and is centered at $(a, b, 0)$ and has radius r is

$$(x - a)^2 + (y - b)^2 = r^2, \quad (1.32)$$

where the value of the z coordinate is unrestricted. Similar equations can be written when the base circle lies in one of the other coordinate planes. A plane intersects a right circular cylinder in a circle, ellipse, or one or two lines, depending on whether that plane is parallel, oblique⁹, or perpendicular, respectively, to the plane containing C . The intersection of a surface with a plane is called the **trace** of the surface.

⁹i.e. at an angle strictly between 0° and 90° .

The equations of spheres and cylinders are examples of *second-degree equations* in \mathbb{R}^3 , i.e. equations of the form

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Iz + J = 0 \quad (1.33)$$

for some constants A, B, \dots, J . If the above equation is not that of a sphere, cylinder, plane, line or point, then the resulting surface is called a **quadric surface**.

One type of quadric surface is the **ellipsoid**, given by an equation of the form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \quad (1.34)$$

In the case where $a = b = c$, this is just a sphere. In general, an ellipsoid is egg-shaped (think of an ellipse rotated around its major axis). Its traces in the coordinate planes are ellipses.

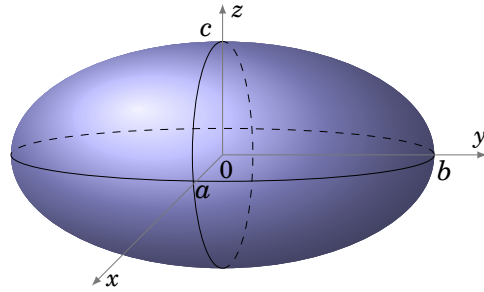


Figure 1.6.4 Ellipsoid

Two other types of quadric surfaces are the **hyperboloid of one sheet**, given by an equation of the form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (1.35)$$

and the **hyperboloid of two sheets**, whose equation has the form:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (1.36)$$

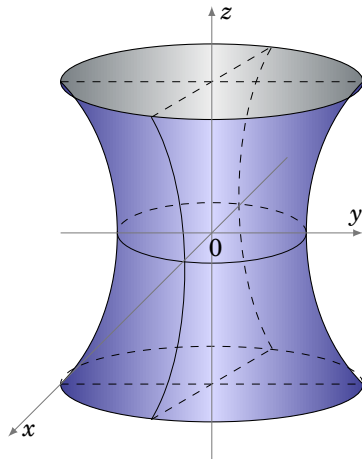


Figure 1.6.5 Hyperboloid of one sheet

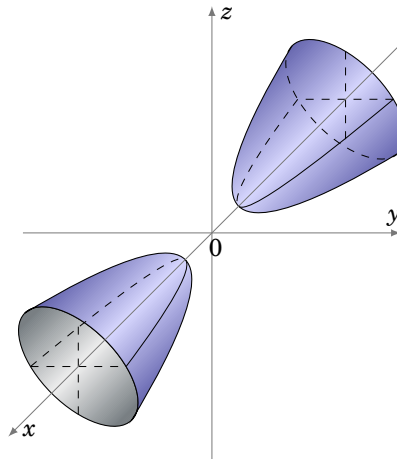


Figure 1.6.6 Hyperboloid of two sheets

For the hyperboloid of one sheet, the trace in any plane parallel to the xy -plane is an ellipse. The traces in the planes parallel to the xz - or yz -planes are hyperbolas (see Figure 1.6.5), except for the special cases $x = \pm a$ and $y = \pm b$; in those planes the traces are pairs of intersecting lines (see Exercise 8).

For the hyperboloid of two sheets, the trace in any plane parallel to the xy - or xz -plane is a hyperbola (see Figure 1.6.6). There is no trace in the yz -plane. In any plane parallel to the yz -plane for which $|x| > |a|$, the trace is an ellipse.

The **elliptic paraboloid** is another type of quadric surface, whose equation has the form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z}{c} \quad (1.37)$$

The traces in planes parallel to the xy -plane are ellipses, though in the xy -plane itself the trace is a single point. The traces in planes parallel to the xz - or yz -planes are parabolas. Figure 1.6.7 shows the case where $c > 0$. When $c < 0$ the surface is turned downward. In the case where $a = b$, the surface is called a *paraboloid of revolution*, which is often used as a reflecting surface, e.g. in vehicle headlights.¹⁰

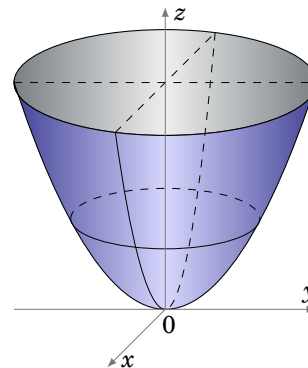


Figure 1.6.7 Paraboloid

A more complicated quadric surface is the **hyperbolic paraboloid**, given by:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c} \quad (1.38)$$

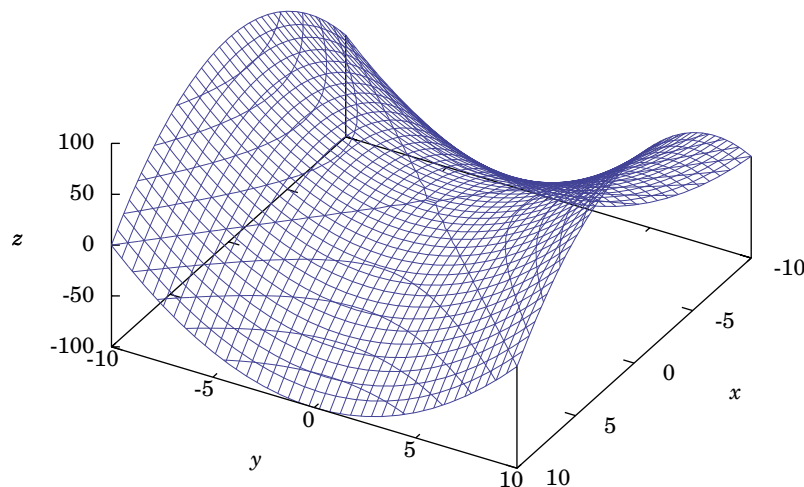


Figure 1.6.8 Hyperbolic paraboloid

¹⁰ For a discussion of this see pp. 157-158 in HECHT.

The hyperbolic paraboloid can be tricky to draw; using graphing software on a computer can make it easier. For example, Figure 1.6.8 was created using the free Gnuplot package (see Appendix C). It shows the graph of the hyperbolic paraboloid $z = y^2 - x^2$, which is the special case where $a = b = 1$ and $c = -1$ in equation (1.38). The mesh lines on the surface are the traces in planes parallel to the coordinate planes. So we see that the traces in planes parallel to the xz -plane are parabolas pointing upward, while the traces in planes parallel to the yz -plane are parabolas pointing downward. Also, notice that the traces in planes parallel to the xy -plane are hyperbolas, though in the xy -plane itself the trace is a pair of intersecting lines through the origin. This is true in general when $c < 0$ in equation (1.38). When $c > 0$, the surface would be similar to that in Figure 1.6.8, only rotated 90° around the z -axis and the nature of the traces in planes parallel to the xz - or yz -planes would be reversed.

The last type of quadric surface that we will consider is the **elliptic cone**, which has an equation of the form:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0 \quad (1.39)$$

The traces in planes parallel to the xy -plane are ellipses, except in the xy -plane itself where the trace is a single point. The traces in planes parallel to the xz - or yz -planes are hyperbolas, except in the xz - and yz -planes themselves where the traces are pairs of intersecting lines.

Notice that every point on the elliptic cone is on a line which lies entirely on the surface; in Figure 1.6.9 these lines all go through the origin. This makes the elliptic cone an example of a *ruled surface*. The cylinder is also a ruled surface.

What may not be as obvious is that both the hyperboloid of one sheet and the hyperbolic paraboloid are ruled surfaces. In fact, on both surfaces there are *two* lines through each point on the surface (see Exercises 11-12). Such surfaces are called *doubly ruled surfaces*, and the pairs of lines are called a *regulus*.

It is clear that for each of the six types of quadric surfaces that we discussed, the surface can be translated away from the origin (e.g. by replacing x^2 by $(x - x_0)^2$ in its equation). It can be proved¹¹ that *every* quadric surface can be translated and/or rotated so that its equation matches one of the six types that we described. For example, $z = 2xy$ is a case of equation (1.33) with “mixed” variables, e.g. with $D \neq 0$ so that we get an xy term. This equation does not match any of the types we considered. However, by rotating the x - and y -axes by 45° in the xy -plane by means of the coordinate transformation $x = (x' - y')/\sqrt{2}$, $y = (x' + y')/\sqrt{2}$, $z = z'$, then $z = 2xy$ becomes the hyperbolic paraboloid $z' = (x')^2 - (y')^2$ in the (x', y', z') coordinate system. That is, $z = 2xy$ is a hyperbolic paraboloid as in equation (1.38), but rotated 45° in the xy -plane.

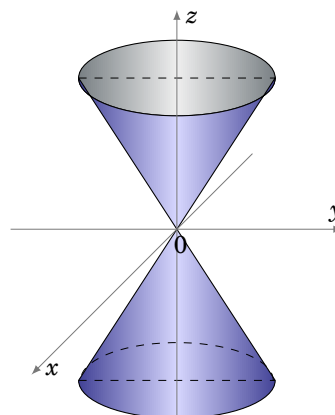


Figure 1.6.9 Elliptic cone

¹¹See Ch. 7 in POGORELOV.

Exercises

A

For Exercises 1-4, determine if the given equation describes a sphere. If so, find its radius and center.

1. $x^2 + y^2 + z^2 - 4x - 6y - 10z + 37 = 0$
2. $x^2 + y^2 + z^2 + 2x - 2y - 8z + 19 = 0$
3. $2x^2 + 2y^2 + 2z^2 + 4x + 4y + 4z - 44 = 0$
4. $x^2 + y^2 - z^2 + 12x + 2y - 4z + 32 = 0$
5. Find the point(s) of intersection of the sphere $(x - 3)^2 + (y + 1)^2 + (z - 3)^2 = 9$ and the line $x = -1 + 2t$, $y = -2 - 3t$, $z = 3 + t$.

B

6. Find the intersection of the spheres $x^2 + y^2 + z^2 = 9$ and $(x - 4)^2 + (y + 2)^2 + (z - 4)^2 = 9$.
7. Find the intersection of the sphere $x^2 + y^2 + z^2 = 9$ and the cylinder $x^2 + y^2 = 4$.
8. Find the trace of the hyperboloid of one sheet $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ in the plane $x = a$, and the trace in the plane $y = b$.
9. Find the trace of the hyperbolic paraboloid $\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{z}{c}$ in the xy -plane.

C

10. It can be shown that any four *noncoplanar* points (i.e. points that do not lie in the same plane) determine a sphere.¹² Find the equation of the sphere that passes through the points $(0, 0, 0)$, $(0, 0, 2)$, $(1, -4, 3)$ and $(0, -1, 3)$. (*Hint: Equation (1.31)*)
11. Show that the hyperboloid of one sheet is a doubly ruled surface, i.e. each point on the surface is on two lines lying entirely on the surface. (*Hint: Write equation (1.35) as $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2}$, factor each side. Recall that two planes intersect in a line.*)
12. Show that the hyperbolic paraboloid is a doubly ruled surface. (*Hint: Exercise 11*)
13. Let S be the sphere with radius 1 centered at $(0, 0, 1)$, and let S^* be S without the “north pole” point $(0, 0, 2)$. Let (a, b, c) be an arbitrary point on S^* . Then the line passing through $(0, 0, 2)$ and (a, b, c) intersects the xy -plane at some point $(x, y, 0)$, as in Figure 1.6.10. Find this point $(x, y, 0)$ in terms of a , b and c .
(Note: Every point in the xy -plane can be matched with a point on S^* , and vice versa, in this manner. This method is called *stereographic projection*, which essentially identifies all of \mathbb{R}^2 with a “punctured” sphere.)

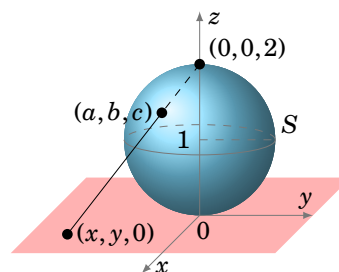


Figure 1.6.10

¹²See WELCHONS and KRICKENBERGER, p. 160, for a proof.

1.7 Curvilinear Coordinates

The Cartesian coordinates of a point (x, y, z) are determined by following straight paths starting from the origin: first along the x -axis, then parallel to the y -axis, then parallel to the z -axis, as in Figure 1.7.1. In *curvilinear coordinate systems*, these paths can be curved. The two types of curvilinear coordinates which we will consider are cylindrical and spherical coordinates. Instead of referencing a point in terms of sides of a rectangular parallelepiped, as with Cartesian coordinates, we will think of the point as lying on a cylinder or sphere. Cylindrical coordinates are often used when there is symmetry around the z -axis; spherical coordinates are useful when there is symmetry about the origin.

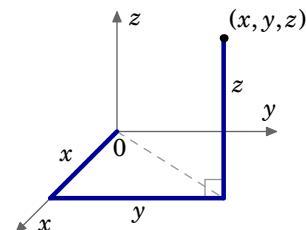


Figure 1.7.1

Let $P = (x, y, z)$ be a point in Cartesian coordinates in \mathbb{R}^3 , and let $P_0 = (x, y, 0)$ be the projection of P upon the xy -plane. Treating (x, y) as a point in \mathbb{R}^2 , let (r, θ) be its polar coordinates (see Figure 1.7.2). Let ρ be the length of the line segment from the origin to P , and let ϕ be the angle between that line segment and the positive z -axis (see Figure 1.7.3). ϕ is called the *zenith angle*. Then the **cylindrical coordinates** (r, θ, z) and the **spherical coordinates** (ρ, θ, ϕ) of $P(x, y, z)$ are defined as follows:¹³

Cylindrical coordinates (r, θ, z) :

$$\begin{aligned} x &= r \cos \theta & r &= \sqrt{x^2 + y^2} \\ y &= r \sin \theta & \theta &= \tan^{-1}\left(\frac{y}{x}\right) \\ z &= z & z &= z \end{aligned}$$

where $0 \leq \theta \leq \pi$ if $y \geq 0$ and $\pi < \theta < 2\pi$ if $y < 0$

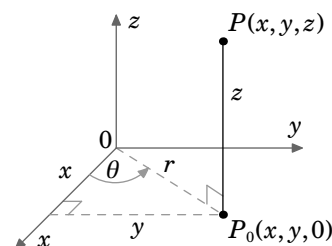


Figure 1.7.2

Cylindrical coordinates

Spherical coordinates (ρ, θ, ϕ) :

$$\begin{aligned} x &= \rho \sin \phi \cos \theta & \rho &= \sqrt{x^2 + y^2 + z^2} \\ y &= \rho \sin \phi \sin \theta & \theta &= \tan^{-1}\left(\frac{y}{x}\right) \\ z &= \rho \cos \phi & \phi &= \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \end{aligned}$$

where $0 \leq \theta \leq \pi$ if $y \geq 0$ and $\pi < \theta < 2\pi$ if $y < 0$

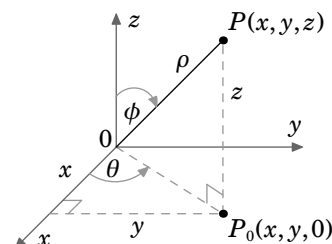


Figure 1.7.3

Spherical coordinates

Both θ and ϕ are measured in radians. Note that $r \geq 0$, $0 \leq \theta < 2\pi$, $\rho \geq 0$ and $0 \leq \phi \leq \pi$. Also, θ is undefined when $(x, y) = (0, 0)$, and ϕ is undefined when $(x, y, z) = (0, 0, 0)$.

¹³This “standard” definition of spherical coordinates used by mathematicians results in a left-handed system. For this reason, physicists usually switch the definitions of θ and ϕ to make (ρ, θ, ϕ) a right-handed system.

Example 1.31. Convert the point $(-2, -2, 1)$ from Cartesian coordinates to (a) cylindrical and (b) spherical coordinates.

Solution: (a) $r = \sqrt{(-2)^2 + (-2)^2} = 2\sqrt{2}$, $\theta = \tan^{-1}\left(\frac{-2}{-2}\right) = \tan^{-1}(1) = \frac{5\pi}{4}$, since $y = -2 < 0$.

$\therefore (r, \theta, z) = (2\sqrt{2}, \frac{5\pi}{4}, 1)$

(b) $\rho = \sqrt{(-2)^2 + (-2)^2 + 1^2} = \sqrt{9} = 3$, $\phi = \cos^{-1}\left(\frac{1}{3}\right) \approx 1.23$ radians.

$\therefore (\rho, \theta, \phi) = (3, \frac{5\pi}{4}, 1.23)$

For cylindrical coordinates (r, θ, z) , and constants r_0 , θ_0 and z_0 , we see from Figure 1.7.4 that the surface $r = r_0$ is a cylinder of radius r_0 centered along the z -axis, the surface $\theta = \theta_0$ is a half-plane emanating from the z -axis, and the surface $z = z_0$ is a plane parallel to the xy -plane.

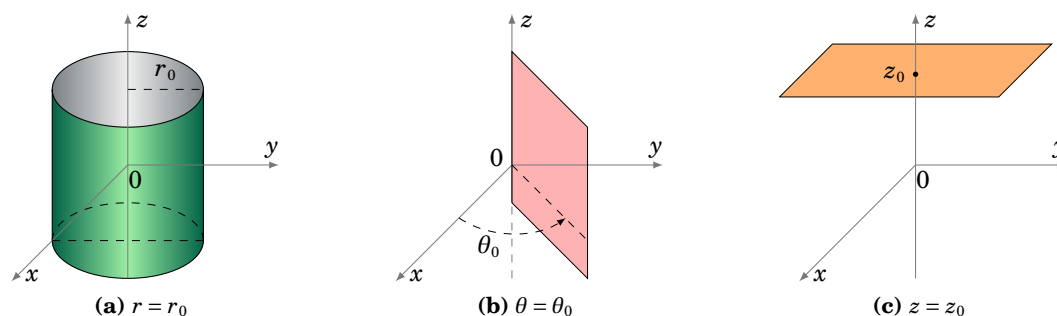


Figure 1.7.4 Cylindrical coordinate surfaces

For spherical coordinates (ρ, θ, ϕ) , and constants ρ_0 , θ_0 and ϕ_0 , we see from Figure 1.7.5 that the surface $\rho = \rho_0$ is a sphere of radius ρ_0 centered at the origin, the surface $\theta = \theta_0$ is a half-plane emanating from the z -axis, and the surface $\phi = \phi_0$ is a circular cone whose vertex is at the origin.

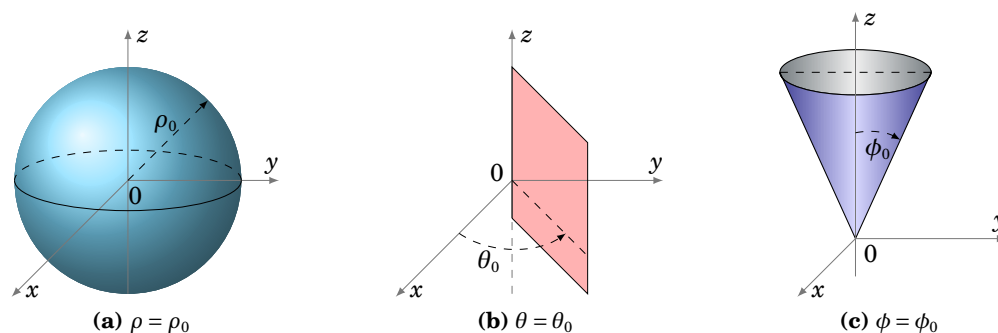


Figure 1.7.5 Spherical coordinate surfaces

Figures 1.7.4(a) and 1.7.5(a) show how these coordinate systems got their names.

Sometimes the equation of a surface in Cartesian coordinates can be transformed into a simpler equation in some other coordinate system, as in the following example.

Example 1.32. Write the equation of the cylinder $x^2 + y^2 = 4$ in cylindrical coordinates.

Solution: Since $r = \sqrt{x^2 + y^2}$, then the equation in cylindrical coordinates is $r = 2$.

Using spherical coordinates to write the equation of a sphere does not necessarily make the equation simpler, if the sphere is not centered at the origin.

Example 1.33. Write the equation $(x - 2)^2 + (y - 1)^2 + z^2 = 9$ in spherical coordinates.

Solution: Multiplying the equation out gives

$$\begin{aligned} x^2 + y^2 + z^2 - 4x - 2y + 5 &= 9, \text{ so we get} \\ \rho^2 - 4\rho \sin \phi \cos \theta - 2\rho \sin \phi \sin \theta - 4 &= 0, \text{ or} \\ \rho^2 - 2 \sin \phi (2 \cos \theta - \sin \theta) \rho - 4 &= 0 \end{aligned}$$

after combining terms. Note that this actually makes it more difficult to figure out what the surface is, as opposed to the Cartesian equation where you could immediately identify the surface as a sphere of radius 3 centered at $(2, 1, 0)$.

Example 1.34. Describe the surface given by $\theta = z$ in cylindrical coordinates.

Solution: This surface is called a *helicoid*. As the (vertical) z coordinate increases, so does the angle θ , while the radius r is unrestricted. So this sweeps out a (ruled!) surface shaped like a spiral staircase, where the spiral has an infinite radius. Figure 1.7.6 shows a section of this surface restricted to $0 \leq z \leq 4\pi$ and $0 \leq r \leq 2$.

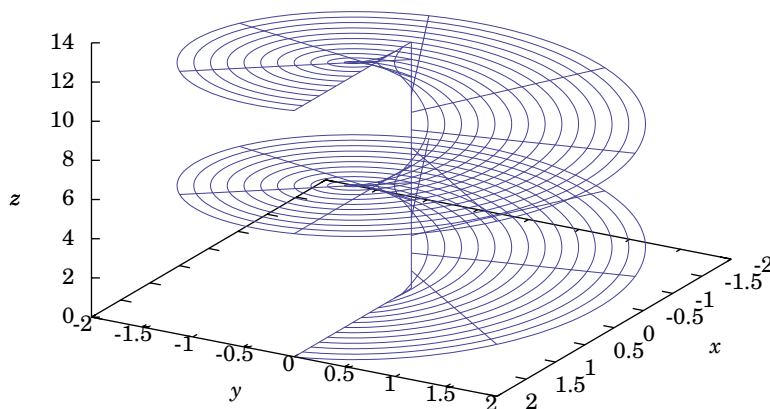


Figure 1.7.6 Helicoid $\theta = z$

Exercises

A

For Exercises 1-4, find the (a) cylindrical and (b) spherical coordinates of the point whose Cartesian coordinates are given.

1. $(2, 2\sqrt{3}, -1)$ 2. $(-5, 5, 6)$ 3. $(\sqrt{21}, -\sqrt{7}, 0)$ 4. $(0, \sqrt{2}, 2)$

For Exercises 5-7, write the given equation in (a) cylindrical and (b) spherical coordinates.

5. $x^2 + y^2 + z^2 = 25$ 6. $x^2 + y^2 = 2y$ 7. $x^2 + y^2 + 9z^2 = 36$

B

8. Describe the intersection of the surfaces whose equations in spherical coordinates are $\theta = \frac{\pi}{2}$ and $\phi = \frac{\pi}{4}$.
9. Show that for $a \neq 0$, the equation $\rho = 2a \sin \phi \cos \theta$ in spherical coordinates describes a sphere centered at $(a, 0, 0)$ with radius $|a|$.

C

10. Let $P = (a, \theta, \phi)$ be a point in spherical coordinates, with $a > 0$ and $0 < \phi < \pi$. Then P lies on the sphere $\rho = a$. Since $0 < \phi < \pi$, the line segment from the origin to P can be extended to intersect the cylinder given by $r = a$ (in cylindrical coordinates). Find the cylindrical coordinates of that point of intersection.
11. Let P_1 and P_2 be points whose spherical coordinates are $(\rho_1, \theta_1, \phi_1)$ and $(\rho_2, \theta_2, \phi_2)$, respectively. Let \mathbf{v}_1 be the vector from the origin to P_1 , and let \mathbf{v}_2 be the vector from the origin to P_2 . For the angle γ between \mathbf{v}_1 and \mathbf{v}_2 , show that

$$\cos \gamma = \cos \phi_1 \cos \phi_2 + \sin \phi_1 \sin \phi_2 \cos(\theta_2 - \theta_1).$$

This formula is used in electrodynamics to prove the addition theorem for spherical harmonics, which provides a general expression for the electrostatic potential at a point due to a unit charge. See pp. 100-102 in JACKSON.

12. Show that the distance d between the points P_1 and P_2 with cylindrical coordinates (r_1, θ_1, z_1) and (r_2, θ_2, z_2) , respectively, is

$$d = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_2 - \theta_1) + (z_2 - z_1)^2}.$$

13. Show that the distance d between the points P_1 and P_2 with spherical coordinates $(\rho_1, \theta_1, \phi_1)$ and $(\rho_2, \theta_2, \phi_2)$, respectively, is

$$d = \sqrt{\rho_1^2 + \rho_2^2 - 2\rho_1 \rho_2 [\sin \phi_1 \sin \phi_2 \cos(\theta_2 - \theta_1) + \cos \phi_1 \cos \phi_2]}.$$

1.8 Vector-Valued Functions

Now that we are familiar with vectors and their operations, we can begin discussing functions whose values are vectors.

Definition 1.10. A **vector-valued function of a real variable** is a rule that associates a vector $\mathbf{f}(t)$ with a real number t , where t is in some subset D of \mathbb{R}^1 (called the **domain** of \mathbf{f}). We write $\mathbf{f}: D \rightarrow \mathbb{R}^3$ to denote that \mathbf{f} is a mapping of D into \mathbb{R}^3 .

For example, $\mathbf{f}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$ is a vector-valued function in \mathbb{R}^3 , defined for all real numbers t . We would write $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^3$. At $t = 1$ the value of the function is the vector $\mathbf{i} + \mathbf{j} + \mathbf{k}$, which in Cartesian coordinates has the terminal point $(1, 1, 1)$.

A vector-valued function of a real variable can be written in component form as

$$\mathbf{f}(t) = f_1(t)\mathbf{i} + f_2(t)\mathbf{j} + f_3(t)\mathbf{k}$$

or in the form

$$\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$$

for some real-valued functions $f_1(t)$, $f_2(t)$, $f_3(t)$, called the *component functions* of \mathbf{f} . The first form is often used when emphasizing that $\mathbf{f}(t)$ is a vector, and the second form is useful when considering just the terminal points of the vectors. By identifying vectors with their terminal points, a curve in space can be written as a vector-valued function.

Example 1.35. Define $\mathbf{f}: \mathbb{R} \rightarrow \mathbb{R}^3$ by $\mathbf{f}(t) = (\cos t, \sin t, t)$.

This is the equation of a *helix* (see Figure 1.8.1). As the value of t increases, the terminal points of $\mathbf{f}(t)$ trace out a curve spiraling upward. For each t , the x - and y -coordinates of $\mathbf{f}(t)$ are $x = \cos t$ and $y = \sin t$, so

$$x^2 + y^2 = \cos^2 t + \sin^2 t = 1.$$

Thus, the curve lies on the surface of the right circular cylinder $x^2 + y^2 = 1$.

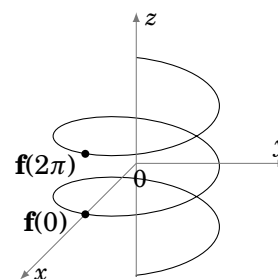


Figure 1.8.1

It may help to think of vector-valued functions of a real variable in \mathbb{R}^3 as a generalization of the parametric functions in \mathbb{R}^2 which you learned about in single-variable calculus. Much of the theory of real-valued functions of a single real variable can be applied to vector-valued functions of a real variable. Since each of the three component functions are real-valued, it will sometimes be the case that results from single-variable calculus can simply be applied to each of the component functions to yield a similar result for the vector-valued function. However, there are times when such generalizations do not hold (see Exercise 13). The concept of a limit, though, can be extended naturally to vector-valued functions, as in the following definition.

Definition 1.11. Let $\mathbf{f}(t)$ be a vector-valued function, let a be a real number and let \mathbf{c} be a vector. Then we say that the **limit** of $\mathbf{f}(t)$ as t approaches a equals \mathbf{c} , written as $\lim_{t \rightarrow a} \mathbf{f}(t) = \mathbf{c}$, if $\lim_{t \rightarrow a} \|\mathbf{f}(t) - \mathbf{c}\| = 0$. If $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$, then

$$\lim_{t \rightarrow a} \mathbf{f}(t) = \left(\lim_{t \rightarrow a} f_1(t), \lim_{t \rightarrow a} f_2(t), \lim_{t \rightarrow a} f_3(t) \right)$$

provided that all three limits on the right side exist.

The above definition shows that continuity and the derivative of vector-valued functions can also be defined in terms of its component functions.

Definition 1.12. Let $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ be a vector-valued function, and let a be a real number in its domain. Then $\mathbf{f}(t)$ is **continuous** at a if $\lim_{t \rightarrow a} \mathbf{f}(t) = \mathbf{f}(a)$. Equivalently, $\mathbf{f}(t)$ is continuous at a if and only if $f_1(t)$, $f_2(t)$, and $f_3(t)$ are continuous at a .

The **derivative** of $\mathbf{f}(t)$ at a , denoted by $\mathbf{f}'(a)$ or $\frac{d\mathbf{f}}{dt}(a)$, is the limit

$$\mathbf{f}'(a) = \lim_{h \rightarrow 0} \frac{\mathbf{f}(a+h) - \mathbf{f}(a)}{h}$$

if that limit exists. Equivalently, $\mathbf{f}'(a) = (f_1'(a), f_2'(a), f_3'(a))$, if the component derivatives exist. We say that $\mathbf{f}(t)$ is **differentiable** at a if $\mathbf{f}'(a)$ exists.

Recall that the derivative of a real-valued function of a single variable is a real number, representing the slope of the tangent line to the graph of the function at a point. Similarly, the derivative of a vector-valued function is a **tangent vector** to the curve in space which the function represents, and it lies on the *tangent line* to the curve (see Figure 1.8.2).

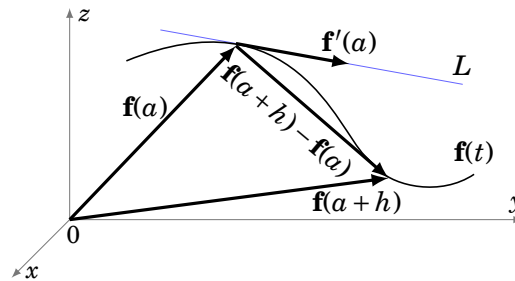


Figure 1.8.2 Tangent vector $\mathbf{f}'(a)$ and tangent line $L = \mathbf{f}(a) + s\mathbf{f}'(a)$

Example 1.36. Let $\mathbf{f}(t) = (\cos t, \sin t, t)$. Then $\mathbf{f}'(t) = (-\sin t, \cos t, 1)$ for all t . The tangent line L to the curve at $\mathbf{f}(2\pi) = (1, 0, 2\pi)$ is $L = \mathbf{f}(2\pi) + s\mathbf{f}'(2\pi) = (1, 0, 2\pi) + s(0, 1, 1)$, or in parametric form: $x = 1$, $y = s$, $z = 2\pi + s$ for $-\infty < s < \infty$.

A **scalar function** is a real-valued function. Note that if $u(t)$ is a scalar function and $\mathbf{f}(t)$ is a vector-valued function, then their product, defined by $(u\mathbf{f})(t) = u(t)\mathbf{f}(t)$ for all t , is a vector-valued function (since the product of a scalar with a vector is a vector).

The basic properties of derivatives of vector-valued functions are summarized in the following theorem.

Theorem 1.20. Let $\mathbf{f}(t)$ and $\mathbf{g}(t)$ be differentiable vector-valued functions, let $u(t)$ be a differentiable scalar function, let k be a scalar, and let \mathbf{c} be a constant vector. Then

$$(a) \frac{d}{dt}(\mathbf{c}) = \mathbf{0}$$

$$(b) \frac{d}{dt}(k\mathbf{f}) = k \frac{d\mathbf{f}}{dt}$$

$$(c) \frac{d}{dt}(\mathbf{f} + \mathbf{g}) = \frac{d\mathbf{f}}{dt} + \frac{d\mathbf{g}}{dt}$$

$$(d) \frac{d}{dt}(\mathbf{f} - \mathbf{g}) = \frac{d\mathbf{f}}{dt} - \frac{d\mathbf{g}}{dt}$$

$$(e) \frac{d}{dt}(u\mathbf{f}) = \frac{du}{dt}\mathbf{f} + u \frac{d\mathbf{f}}{dt}$$

$$(f) \frac{d}{dt}(\mathbf{f} \cdot \mathbf{g}) = \frac{d\mathbf{f}}{dt} \cdot \mathbf{g} + \mathbf{f} \cdot \frac{d\mathbf{g}}{dt}$$

$$(g) \frac{d}{dt}(\mathbf{f} \times \mathbf{g}) = \frac{d\mathbf{f}}{dt} \times \mathbf{g} + \mathbf{f} \times \frac{d\mathbf{g}}{dt}$$

Proof: The proofs of parts (a)-(e) follow easily by differentiating the component functions and using the rules for derivatives from single-variable calculus. We will prove part (f), and leave the proof of part (g) as an exercise for the reader.

(f) Write $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ and $\mathbf{g}(t) = (g_1(t), g_2(t), g_3(t))$, where the component functions $f_1(t), f_2(t), f_3(t), g_1(t), g_2(t), g_3(t)$ are all differentiable real-valued functions. Then

$$\begin{aligned} \frac{d}{dt}(\mathbf{f}(t) \cdot \mathbf{g}(t)) &= \frac{d}{dt}(f_1(t)g_1(t) + f_2(t)g_2(t) + f_3(t)g_3(t)) \\ &= \frac{d}{dt}(f_1(t)g_1(t)) + \frac{d}{dt}(f_2(t)g_2(t)) + \frac{d}{dt}(f_3(t)g_3(t)) \\ &= \frac{df_1}{dt}(t)g_1(t) + f_1(t)\frac{dg_1}{dt}(t) + \frac{df_2}{dt}(t)g_2(t) + f_2(t)\frac{dg_2}{dt}(t) + \frac{df_3}{dt}(t)g_3(t) + f_3(t)\frac{dg_3}{dt}(t) \\ &= \left(\frac{df_1}{dt}(t), \frac{df_2}{dt}(t), \frac{df_3}{dt}(t)\right) \cdot (g_1(t), g_2(t), g_3(t)) \\ &\quad + (f_1(t), f_2(t), f_3(t)) \cdot \left(\frac{dg_1}{dt}(t), \frac{dg_2}{dt}(t), \frac{dg_3}{dt}(t)\right) \\ &= \frac{d\mathbf{f}}{dt}(t) \cdot \mathbf{g}(t) + \mathbf{f}(t) \cdot \frac{d\mathbf{g}}{dt}(t) \text{ for all } t. \end{aligned}$$

QED

Example 1.37. Suppose $\mathbf{f}(t)$ is differentiable. Find the derivative of $\|\mathbf{f}(t)\|$.

Solution: Since $\|\mathbf{f}(t)\|$ is a real-valued function of t , then by the Chain Rule for real-valued

functions, we know that $\frac{d}{dt}\|\mathbf{f}(t)\|^2 = 2\|\mathbf{f}(t)\| \frac{d}{dt}\|\mathbf{f}(t)\|$.

But $\|\mathbf{f}(t)\|^2 = \mathbf{f}(t) \cdot \mathbf{f}(t)$, so $\frac{d}{dt}\|\mathbf{f}(t)\|^2 = \frac{d}{dt}(\mathbf{f}(t) \cdot \mathbf{f}(t))$. Hence, we have

$$\begin{aligned} 2\|\mathbf{f}(t)\| \frac{d}{dt}\|\mathbf{f}(t)\| &= \frac{d}{dt}(\mathbf{f}(t) \cdot \mathbf{f}(t)) = \mathbf{f}'(t) \cdot \mathbf{f}(t) + \mathbf{f}(t) \cdot \mathbf{f}'(t) \text{ by Theorem 1.20(f), so} \\ &= 2\mathbf{f}'(t) \cdot \mathbf{f}(t), \text{ so if } \|\mathbf{f}(t)\| \neq 0 \text{ then} \\ \frac{d}{dt}\|\mathbf{f}(t)\| &= \frac{\mathbf{f}'(t) \cdot \mathbf{f}(t)}{\|\mathbf{f}(t)\|}. \end{aligned}$$

We know that $\|\mathbf{f}(t)\|$ is constant if and only if $\frac{d}{dt}\|\mathbf{f}(t)\| = 0$ for all t . Also, $\mathbf{f}(t) \perp \mathbf{f}'(t)$ if and only if $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$. Thus, the above example shows this important fact:

If $\|\mathbf{f}(t)\| \neq 0$, then $\|\mathbf{f}(t)\|$ is constant if and only if $\mathbf{f}(t) \perp \mathbf{f}'(t)$ for all t .

This means that if a curve lies completely on a sphere (or circle) centered at the origin, then the tangent vector $\mathbf{f}'(t)$ is always perpendicular to the *position vector* $\mathbf{f}(t)$.

Example 1.38. The *spherical spiral* $\mathbf{f}(t) = \left(\frac{\cos t}{\sqrt{1+a^2t^2}}, \frac{\sin t}{\sqrt{1+a^2t^2}}, \frac{-at}{\sqrt{1+a^2t^2}} \right)$, for $a \neq 0$.

Figure 1.8.3 shows the graph of the curve when $a = 0.2$. In the exercises, the reader will be asked to show that this curve lies on the sphere $x^2 + y^2 + z^2 = 1$ and to verify directly that $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$ for all t .

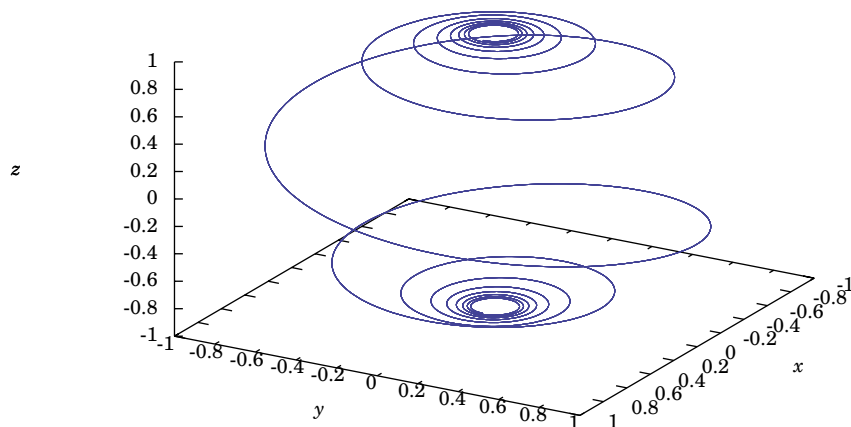


Figure 1.8.3 Spherical spiral with $a = 0.2$

Just as in single-variable calculus, higher-order derivatives of vector-valued functions are obtained by repeatedly differentiating the (first) derivative of the function:

$$\mathbf{f}''(t) = \frac{d}{dt}\mathbf{f}'(t), \quad \mathbf{f}'''(t) = \frac{d}{dt}\mathbf{f}''(t), \quad \dots, \quad \frac{d^n \mathbf{f}}{dt^n} = \frac{d}{dt}\left(\frac{d^{n-1} \mathbf{f}}{dt^{n-1}}\right) \quad (\text{for } n = 2, 3, 4, \dots)$$

We can use vector-valued functions to represent physical quantities, such as velocity, acceleration, force, momentum, etc. For example, let the real variable t represent time elapsed from some initial time ($t = 0$), and suppose that an object of constant mass m is subjected to some force so that it moves in space, with its position (x, y, z) at time t a function of t . That is, $x = x(t)$, $y = y(t)$, $z = z(t)$ for some real-valued functions $x(t)$, $y(t)$, $z(t)$. Call $\mathbf{r}(t) = (x(t), y(t), z(t))$ the **position vector** of the object. We can define various physical quantities associated with the object as follows:¹⁴

$$\text{position: } \mathbf{r}(t) = (x(t), y(t), z(t))$$

$$\begin{aligned} \text{velocity: } \mathbf{v}(t) &= \dot{\mathbf{r}}(t) = \mathbf{r}'(t) = \frac{d\mathbf{r}}{dt} \\ &= (x'(t), y'(t), z'(t)) \end{aligned}$$

$$\begin{aligned} \text{acceleration: } \mathbf{a}(t) &= \dot{\mathbf{v}}(t) = \mathbf{v}'(t) = \frac{d\mathbf{v}}{dt} \\ &= \ddot{\mathbf{r}}(t) = \mathbf{r}''(t) = \frac{d^2 \mathbf{r}}{dt^2} \\ &= (x''(t), y''(t), z''(t)) \end{aligned}$$

$$\text{momentum: } \mathbf{p}(t) = m\mathbf{v}(t)$$

$$\text{force: } \mathbf{F}(t) = \dot{\mathbf{p}}(t) = \mathbf{p}'(t) = \frac{d\mathbf{p}}{dt} \quad (\text{Newton's Second Law of Motion})$$

The magnitude $\|\mathbf{v}(t)\|$ of the velocity vector is called the *speed* of the object. Note that since the mass m is a constant, the force equation becomes the familiar $\mathbf{F}(t) = m\mathbf{a}(t)$.

Example 1.39. Let $\mathbf{r}(t) = (5 \cos t, 3 \sin t, 4 \sin t)$ be the position vector of an object at time $t \geq 0$. Find its (a) velocity and (b) acceleration vectors.

Solution: (a) $\mathbf{v}(t) = \dot{\mathbf{r}}(t) = (-5 \sin t, 3 \cos t, 4 \cos t)$

(b) $\mathbf{a}(t) = \dot{\mathbf{v}}(t) = (-5 \cos t, -3 \sin t, -4 \sin t)$

Note that $\|\mathbf{r}(t)\| = \sqrt{25 \cos^2 t + 25 \sin^2 t} = 5$ for all t , so by Example 1.37 we know that $\mathbf{r}(t) \cdot \dot{\mathbf{r}}(t) = 0$ for all t (which we can verify from part (a)). In fact, $\|\mathbf{v}(t)\| = 5$ for all t also. And not only does $\mathbf{r}(t)$ lie on the sphere of radius 5 centered at the origin, but perhaps not so obvious is that it lies completely within a *circle* of radius 5 centered at the origin. Also, note that $\mathbf{a}(t) = -\mathbf{r}(t)$. It turns out (see Exercise 16) that whenever an object moves in a circle with constant speed, the acceleration vector will point in the opposite direction of the position vector (i.e. towards the center of the circle).

¹⁴We will often use the older dot notation for derivatives when physics is involved.

Recall from Section 1.5 that if $\mathbf{r}_1, \mathbf{r}_2$ are position vectors to distinct points then $\mathbf{r}_1 + t(\mathbf{r}_2 - \mathbf{r}_1)$ represents a line through those two points as t varies over all real numbers. That vector sum can be written as $(1 - t)\mathbf{r}_1 + t\mathbf{r}_2$. So the function $\mathbf{l}(t) = (1 - t)\mathbf{r}_1 + t\mathbf{r}_2$ is a line through the terminal points of \mathbf{r}_1 and \mathbf{r}_2 , and when t is restricted to the interval $[0, 1]$ it is the line segment between the points, with $\mathbf{l}(0) = \mathbf{r}_1$ and $\mathbf{l}(1) = \mathbf{r}_2$.

In general, a function of the form $\mathbf{f}(t) = (a_1t + b_1, a_2t + b_2, a_3t + b_3)$ represents a line in \mathbb{R}^3 . A function of the form $\mathbf{f}(t) = (a_1t^2 + b_1t + c_1, a_2t^2 + b_2t + c_2, a_3t^2 + b_3t + c_3)$ represents a (possibly degenerate) parabola in \mathbb{R}^3 .

Example 1.40. *Bézier curves* are used in Computer Aided Design (CAD) to approximate the shape of a polygonal path in space (called the *Bézier polygon* or *control polygon*). For instance, given three points (or position vectors) $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$ in \mathbb{R}^3 , define

$$\begin{aligned}\mathbf{b}_0^1(t) &= (1 - t)\mathbf{b}_0 + t\mathbf{b}_1 \\ \mathbf{b}_1^1(t) &= (1 - t)\mathbf{b}_1 + t\mathbf{b}_2 \\ \mathbf{b}_0^2(t) &= (1 - t)\mathbf{b}_0^1(t) + t\mathbf{b}_1^1(t) \\ &= (1 - t)^2\mathbf{b}_0 + 2t(1 - t)\mathbf{b}_1 + t^2\mathbf{b}_2\end{aligned}$$

for all real t . For t in the interval $[0, 1]$, we see that $\mathbf{b}_0^1(t)$ is the line segment between \mathbf{b}_0 and \mathbf{b}_1 , and $\mathbf{b}_1^1(t)$ is the line segment between \mathbf{b}_1 and \mathbf{b}_2 . The function $\mathbf{b}_0^2(t)$ is the Bézier curve for the points $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$. Note from the last formula that the curve is a parabola that goes through \mathbf{b}_0 (when $t = 0$) and \mathbf{b}_2 (when $t = 1$).

As an example, let $\mathbf{b}_0 = (0, 0, 0)$, $\mathbf{b}_1 = (1, 2, 3)$, and $\mathbf{b}_2 = (4, 5, 2)$. Then the explicit formula for the Bézier curve is $\mathbf{b}_0^2(t) = (2t + 2t^2, 4t + t^2, 6t - 4t^2)$, as shown in Figure 1.8.4, where the line segments are $\mathbf{b}_0^1(t)$ and $\mathbf{b}_1^1(t)$, and the curve is $\mathbf{b}_0^2(t)$.

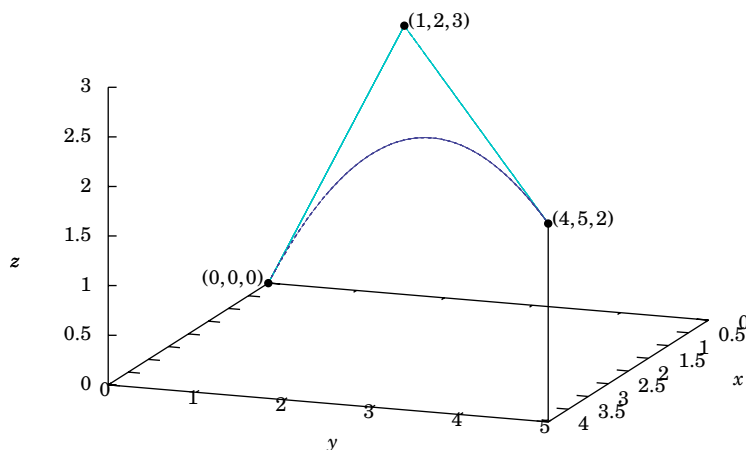


Figure 1.8.4 Bézier curve approximation for three points

In general, the polygonal path determined by $n \geq 3$ noncollinear points in \mathbb{R}^3 can be used to define the Bézier curve recursively by a process called *repeated linear interpolation*. This curve will be a vector-valued function whose components are polynomials of degree $n - 1$, and its formula is given by *de Casteljau's algorithm*.¹⁵ In the exercises, the reader will be given the algorithm for the case of $n = 4$ points and asked to write the explicit formula for the Bézier curve for the four points shown in Figure 1.8.5.

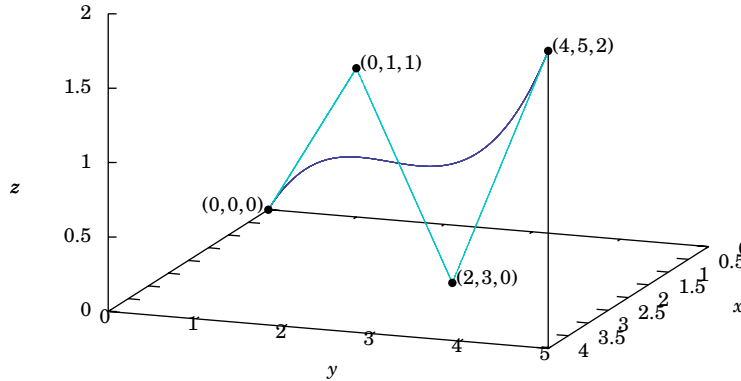


Figure 1.8.5 Bézier curve approximation for four points

Exercises

A

For Exercises 1-4, calculate $\mathbf{f}'(t)$ and find the tangent line at $\mathbf{f}(0)$.

1. $\mathbf{f}(t) = (t + 1, t^2 + 1, t^3 + 1)$
2. $\mathbf{f}(t) = (e^t + 1, e^{2t} + 1, e^{t^2} + 1)$
3. $\mathbf{f}(t) = (\cos 2t, \sin 2t, t)$
4. $\mathbf{f}(t) = (\sin 2t, 2\sin^2 t, 2\cos t)$

For Exercises 5-6, find the velocity $\mathbf{v}(t)$ and acceleration $\mathbf{a}(t)$ of an object with the given position vector $\mathbf{r}(t)$.

5. $\mathbf{r}(t) = (t, t - \sin t, 1 - \cos t)$
6. $\mathbf{r}(t) = (3 \cos t, 2 \sin t, 1)$

B

7. Let $\mathbf{f}(t) = \left(\frac{\cos t}{\sqrt{1+a^2t^2}}, \frac{\sin t}{\sqrt{1+a^2t^2}}, \frac{-at}{\sqrt{1+a^2t^2}} \right)$, with $a \neq 0$.

- (a) Show that $\|\mathbf{f}(t)\| = 1$ for all t .
- (b) Show directly that $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$ for all t .

8. If $\mathbf{f}'(t) = \mathbf{0}$ for all t in some interval (a, b) , show that $\mathbf{f}(t)$ is a constant vector in (a, b) .

¹⁵See pp. 27-30 in FARIN.

9. For a constant vector $\mathbf{c} \neq \mathbf{0}$, the function $\mathbf{f}(t) = t\mathbf{c}$ represents a line parallel to \mathbf{c} .
- (a) What kind of curve does $\mathbf{g}(t) = t^3\mathbf{c}$ represent? Explain.
- (b) What kind of curve does $\mathbf{h}(t) = e^t\mathbf{c}$ represent? Explain.
- (c) Compare $\mathbf{f}'(0)$ and $\mathbf{g}'(0)$. Given your answer to part (a), how do you explain the difference in the two derivatives?
10. Show that $\frac{d}{dt}\left(\mathbf{f} \times \frac{d\mathbf{f}}{dt}\right) = \mathbf{f} \times \frac{d^2\mathbf{f}}{dt^2}$.
11. Let a particle of (constant) mass m have position vector $\mathbf{r}(t)$, velocity $\mathbf{v}(t)$, acceleration $\mathbf{a}(t)$ and momentum $\mathbf{p}(t)$ at time t . The *angular momentum* $\mathbf{L}(t)$ of the particle with respect to the origin at time t is defined as $\mathbf{L}(t) = \mathbf{r}(t) \times \mathbf{p}(t)$. If $\mathbf{F}(t)$ is the force acting on the particle at time t , then define the *torque* $\mathbf{N}(t)$ acting on the particle with respect to the origin as $\mathbf{N}(t) = \mathbf{r}(t) \times \mathbf{F}(t)$. Show that $\mathbf{L}'(t) = \mathbf{N}(t)$.
12. Show that $\frac{d}{dt}(\mathbf{f} \cdot (\mathbf{g} \times \mathbf{h})) = \frac{d\mathbf{f}}{dt} \cdot (\mathbf{g} \times \mathbf{h}) + \mathbf{f} \cdot \left(\frac{d\mathbf{g}}{dt} \times \mathbf{h}\right) + \mathbf{f} \cdot \left(\mathbf{g} \times \frac{d\mathbf{h}}{dt}\right)$.
13. The Mean Value Theorem does not hold for vector-valued functions: Show that for $\mathbf{f}(t) = (\cos t, \sin t, t)$, there is no t in the interval $(0, 2\pi)$ such that

$$\mathbf{f}'(t) = \frac{\mathbf{f}(2\pi) - \mathbf{f}(0)}{2\pi - 0}.$$

C

14. The Bézier curve $\mathbf{b}_0^3(t)$ for four noncollinear points $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ in \mathbb{R}^3 is defined by the following algorithm (going from the left column to the right):
- $$\begin{aligned} \mathbf{b}_0^1(t) &= (1-t)\mathbf{b}_0 + t\mathbf{b}_1 & \mathbf{b}_0^2(t) &= (1-t)\mathbf{b}_0^1(t) + t\mathbf{b}_1^1(t) & \mathbf{b}_0^3(t) &= (1-t)\mathbf{b}_0^2(t) + t\mathbf{b}_1^2(t) \\ \mathbf{b}_1^1(t) &= (1-t)\mathbf{b}_1 + t\mathbf{b}_2 & \mathbf{b}_1^2(t) &= (1-t)\mathbf{b}_1^1(t) + t\mathbf{b}_2^1(t) \\ \mathbf{b}_2^1(t) &= (1-t)\mathbf{b}_2 + t\mathbf{b}_3 \end{aligned}$$
- (a) Show that $\mathbf{b}_0^3(t) = (1-t)^3\mathbf{b}_0 + 3t(1-t)^2\mathbf{b}_1 + 3t^2(1-t)\mathbf{b}_2 + t^3\mathbf{b}_3$.
- (b) Write the explicit formula (as in Example 1.40) for the Bézier curve for the points $\mathbf{b}_0 = (0, 0, 0)$, $\mathbf{b}_1 = (0, 1, 1)$, $\mathbf{b}_2 = (2, 3, 0)$, $\mathbf{b}_3 = (4, 5, 2)$.
15. Let $\mathbf{r}(t)$ be the position vector for a particle moving in \mathbb{R}^3 . Show that
- $$\frac{d}{dt}(\mathbf{r} \times (\mathbf{v} \times \mathbf{r})) = \|\mathbf{r}\|^2 \mathbf{a} + (\mathbf{r} \cdot \mathbf{v})\mathbf{v} - (\|\mathbf{v}\|^2 + \mathbf{r} \cdot \mathbf{a})\mathbf{r}.$$
16. Let $\mathbf{r}(t)$ be the position vector in \mathbb{R}^3 for a particle that moves with constant speed $c > 0$ in a circle of radius $a > 0$ in the xy -plane. Show that $\mathbf{a}(t)$ points in the opposite direction as $\mathbf{r}(t)$ for all t . (Hint: Use Example 1.37 to show that $\mathbf{r}(t) \perp \mathbf{v}(t)$ and $\mathbf{a}(t) \perp \mathbf{v}(t)$, and hence $\mathbf{a}(t) \parallel \mathbf{r}(t)$.)
17. Prove Theorem 1.20(g).

1.9 Arc Length

Let $\mathbf{r}(t) = (x(t), y(t), z(t))$ be the position vector of an object moving in \mathbb{R}^3 . Since $\|\mathbf{v}(t)\|$ is the speed of the object at time t , it seems natural to define the distance s traveled by the object from time $t = a$ to $t = b$ as the definite integral

$$s = \int_a^b \|\mathbf{v}(t)\| dt = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt, \quad (1.40)$$

which is analogous to the case from single-variable calculus for parametric functions in \mathbb{R}^2 . This is indeed how we will define the distance traveled and, in general, the arc length of a curve in \mathbb{R}^3 .

Definition 1.13. Let $\mathbf{f}(t) = (x(t), y(t), z(t))$ be a curve in \mathbb{R}^3 whose domain includes the interval $[a, b]$. Suppose that in the interval (a, b) the first derivative of each component function $x(t)$, $y(t)$ and $z(t)$ exists and is continuous, and that no section of the curve is repeated. Then the **arc length** L of the curve from $t = a$ to $t = b$ is

$$L = \int_a^b \|\mathbf{f}'(t)\| dt = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt \quad (1.41)$$

A real-valued function whose first derivative is continuous is called *continuously differentiable* (or a \mathcal{C}^1 function), and a function whose derivatives of all orders are continuous is called *smooth* (or a \mathcal{C}^∞ function). All the functions we will consider will be smooth. A *smooth curve* $\mathbf{f}(t)$ is one whose derivative $\mathbf{f}'(t)$ is never the zero vector and whose component functions are all smooth.

Note that we did not *prove* that the formula in the above definition actually gives the length of a section of a curve. A rigorous proof requires dealing with some subtleties, normally glossed over in calculus texts, which are beyond the scope of this book.¹⁶

Example 1.41. Find the length L of the helix $\mathbf{f}(t) = (\cos t, \sin t, t)$ from $t = 0$ to $t = 2\pi$.

Solution: By formula (1.41), we have

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{(-\sin t)^2 + (\cos t)^2 + 1^2} dt = \int_0^{2\pi} \sqrt{\sin^2 t + \cos^2 t + 1} dt = \int_0^{2\pi} \sqrt{2} dt \\ &= \sqrt{2}(2\pi - 0) = 2\sqrt{2}\pi \end{aligned}$$

Similar to the case in \mathbb{R}^2 , if there are values of t in the interval $[a, b]$ where the derivative of a component function is not continuous then it is often possible to partition $[a, b]$ into subintervals where all the component functions are continuously differentiable (except at the endpoints, which can be ignored). The sum of the arc lengths over the subintervals will be the arc length over $[a, b]$.

¹⁶In particular, *Duhamel's principle* is needed. See the proof in TAYLOR and MANN, § 14.2 and § 18.2.

Notice that the curve traced out by the function $\mathbf{f}(t) = (\cos t, \sin t, t)$ from Example 1.41 is also traced out by the function $\mathbf{g}(t) = (\cos 2t, \sin 2t, 2t)$. For example, over the interval $[0, \pi]$, $\mathbf{g}(t)$ traces out the same section of the curve as $\mathbf{f}(t)$ does over the interval $[0, 2\pi]$. Intuitively, this says that $\mathbf{g}(t)$ traces the curve twice as fast as $\mathbf{f}(t)$. This makes sense since, viewing the functions as position vectors and their derivatives as velocity vectors, the speeds of $\mathbf{f}(t)$ and $\mathbf{g}(t)$ are $\|\mathbf{f}'(t)\| = \sqrt{2}$ and $\|\mathbf{g}'(t)\| = 2\sqrt{2}$, respectively. We say that $\mathbf{g}(t)$ and $\mathbf{f}(t)$ are different *parametrizations* of the same curve.

Definition 1.14. Let C be a smooth curve in \mathbb{R}^3 represented by a function $\mathbf{f}(t)$ defined on an interval $[a, b]$, and let $\alpha : [c, d] \rightarrow [a, b]$ be a smooth one-to-one mapping of an interval $[c, d]$ onto $[a, b]$. Then the function $\mathbf{g} : [c, d] \rightarrow \mathbb{R}^3$ defined by $\mathbf{g}(s) = \mathbf{f}(\alpha(s))$ is a **parametrization** of C with **parameter** s . If α is strictly increasing on $[c, d]$ then we say that $\mathbf{g}(s)$ is *equivalent* to $\mathbf{f}(t)$.

$$\begin{array}{ccccc} s & & t & & \mathbf{f}(t) \\ [c, d] & \xrightarrow{\alpha} & [a, b] & \xrightarrow{\mathbf{f}} & \mathbb{R}^3 \\ & \searrow & & \nearrow & \\ & & \mathbf{g}(s) = \mathbf{f}(\alpha(s)) = \mathbf{f}(t) & & \end{array}$$

Note that the differentiability of $\mathbf{g}(s)$ follows from a version of the Chain Rule for vector-valued functions (the proof is left as an exercise):

Theorem 1.21. Chain Rule: If $\mathbf{f}(t)$ is a differentiable vector-valued function of t , and $t = \alpha(s)$ is a differentiable scalar function of s , then $\mathbf{f}(s) = \mathbf{f}(\alpha(s))$ is a differentiable vector-valued function of s , and

$$\frac{d\mathbf{f}}{ds} = \frac{d\mathbf{f}}{dt} \frac{dt}{ds} \quad (1.42)$$

for any s where the composite function $\mathbf{f}(\alpha(s))$ is defined.

Example 1.42. The following are all equivalent parametrizations of the same curve:

$$\begin{aligned} \mathbf{f}(t) &= (\cos t, \sin t, t) \text{ for } t \text{ in } [0, 2\pi] \\ \mathbf{g}(s) &= (\cos 2s, \sin 2s, 2s) \text{ for } s \text{ in } [0, \pi] \\ \mathbf{h}(s) &= (\cos 2\pi s, \sin 2\pi s, 2\pi s) \text{ for } s \text{ in } [0, 1] \end{aligned}$$

To see that $\mathbf{g}(s)$ is equivalent to $\mathbf{f}(t)$, define $\alpha : [0, \pi] \rightarrow [0, 2\pi]$ by $\alpha(s) = 2s$. Then α is smooth, one-to-one, maps $[0, \pi]$ onto $[0, 2\pi]$, and is strictly increasing (since $\alpha'(s) = 2 > 0$ for all s). Likewise, defining $\alpha : [0, 1] \rightarrow [0, 2\pi]$ by $\alpha(s) = 2\pi s$ shows that $\mathbf{h}(s)$ is equivalent to $\mathbf{f}(t)$.

A curve can have many parametrizations, with different speeds, so which one is the best to use? In some situations the **arc length parametrization** can be useful. The idea behind this is to replace the parameter t , for any given smooth parametrization $\mathbf{f}(t)$ defined on $[a, b]$, by the parameter s given by

$$s = s(t) = \int_a^t \|\mathbf{f}'(u)\| du. \quad (1.43)$$

In terms of motion along a curve, s is the distance traveled along the curve after time t has elapsed. So the new parameter will be distance instead of time. There is a natural correspondence between s and t : from a starting point on the curve, the distance traveled along the curve (in one direction) is uniquely determined by the amount of time elapsed, and vice versa.

Since s is the arc length of the curve over the interval $[a, t]$ for each t in $[a, b]$, then it is a function of t . By the Fundamental Theorem of Calculus, its derivative is

$$s'(t) = \frac{ds}{dt} = \frac{d}{dt} \int_a^t \|\mathbf{f}'(u)\| du = \|\mathbf{f}'(t)\| \text{ for all } t \text{ in } [a, b].$$

Since $\mathbf{f}(t)$ is smooth, then $\|\mathbf{f}'(t)\| > 0$ for all t in $[a, b]$. Thus $s'(t) > 0$ and hence $s(t)$ is strictly increasing on the interval $[a, b]$. Recall that this means that s is a one-to-one mapping of the interval $[a, b]$ onto the interval $[s(a), s(b)]$. But we see that

$$s(a) = \int_a^a \|\mathbf{f}'(u)\| du = 0 \quad \text{and} \quad s(b) = \int_a^b \|\mathbf{f}'(u)\| du = L = \text{arc length from } t = a \text{ to } t = b$$

So the function $s : [a, b] \rightarrow [0, L]$ is a one-to-one, differentiable mapping onto the interval $[0, L]$. From single-variable calculus, we know that this means that there exists an inverse function $\alpha : [0, L] \rightarrow [a, b]$ that is differentiable and the inverse of $s : [a, b] \rightarrow [0, L]$. That is, for each t in $[a, b]$ there is a unique s in $[0, L]$ such that $s = s(t)$ and $t = \alpha(s)$. And we know that the derivative of α is

$$\alpha'(s) = \frac{1}{s'(\alpha(s))} = \frac{1}{\|\mathbf{f}'(\alpha(s))\|}$$

So define the arc length parametrization $\mathbf{f} : [0, L] \rightarrow \mathbb{R}^3$ by

$$\mathbf{f}(s) = \mathbf{f}(\alpha(s)) \text{ for all } s \text{ in } [0, L].$$

Then $\mathbf{f}(s)$ is smooth, by the Chain Rule. In fact, $\mathbf{f}(s)$ has *unit speed*:

$$\begin{aligned} \mathbf{f}'(s) &= \mathbf{f}'(\alpha(s)) \alpha'(s) \text{ by the Chain Rule, so} \\ &= \mathbf{f}'(\alpha(s)) \frac{1}{\|\mathbf{f}'(\alpha(s))\|}, \text{ so} \\ \|\mathbf{f}'(s)\| &= 1 \text{ for all } s \text{ in } [0, L]. \end{aligned}$$

So the arc length parametrization traverses the curve at a “normal” rate.

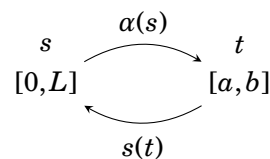


Figure 1.9.1 $t = \alpha(s)$

In practice, parametrizing a curve $\mathbf{f}(t)$ by arc length requires you to evaluate the integral $s = \int_a^t \|\mathbf{f}'(u)\| du$ in some closed form (as a function of t) so that you could then solve for t in terms of s . If that can be done, you would then substitute the expression for t in terms of s (which we called $\alpha(s)$) into the formula for $\mathbf{f}(t)$ to get $\mathbf{f}(s)$.

Example 1.43. Parametrize the helix $\mathbf{f}(t) = (\cos t, \sin t, t)$, for t in $[0, 2\pi]$, by arc length.

Solution: By Example 1.41 and formula (1.43), we have

$$s = \int_0^t \|\mathbf{f}'(u)\| du = \int_0^t \sqrt{2} du = \sqrt{2}t \text{ for all } t \text{ in } [0, 2\pi].$$

So we can solve for t in terms of s : $t = \alpha(s) = \frac{s}{\sqrt{2}}$.

$\therefore \mathbf{f}(s) = \left(\cos \frac{s}{\sqrt{2}}, \sin \frac{s}{\sqrt{2}}, \frac{s}{\sqrt{2}} \right)$ for all s in $[0, 2\sqrt{2}\pi]$. Note that $\|\mathbf{f}'(s)\| = 1$.

Arc length plays an important role when discussing *curvature* and *moving frame fields*, in the field of mathematics known as *differential geometry*.¹⁷ The methods involve using an arc length parametrization, which often leads to an integral that is either difficult or impossible to evaluate in a simple closed form. The simple integral in Example 1.43 is the exception, not the norm. In general, arc length parametrizations are more useful for theoretical purposes than for practical computations.¹⁸ Curvature and moving frame fields can be defined without using arc length, which makes their computation much easier, and these definitions can be shown to be equivalent to those using arc length. We will leave this to the exercises.

The arc length for curves given in other coordinate systems can also be calculated:

Theorem 1.22. Suppose that $r = r(t)$, $\theta = \theta(t)$ and $z = z(t)$ are the cylindrical coordinates of a curve $\mathbf{f}(t)$, for t in $[a, b]$. Then the arc length L of the curve over $[a, b]$ is

$$L = \int_a^b \sqrt{r'(t)^2 + r(t)^2 \theta'(t)^2 + z'(t)^2} dt \quad (1.44)$$

Proof: The Cartesian coordinates $(x(t), y(t), z(t))$ of a point on the curve are given by

$$x(t) = r(t) \cos \theta(t), \quad y(t) = r(t) \sin \theta(t), \quad z(t) = z(t)$$

so differentiating the above expressions for $x(t)$ and $y(t)$ with respect to t gives

$$x'(t) = r'(t) \cos \theta(t) - r(t) \theta'(t) \sin \theta(t), \quad y'(t) = r'(t) \sin \theta(t) + r(t) \theta'(t) \cos \theta(t)$$

¹⁷See O'NEILL for an introduction to elementary differential geometry.

¹⁸For example, the usual parametrizations of Bézier curves, which we discussed in Section 1.8, are polynomial functions in \mathbb{R}^3 . This makes their computation relatively simple, which, in CAD, is desirable. But their arc length parametrizations are not only *not* polynomials, they are in fact usually impossible to calculate at all.

and so

$$\begin{aligned}
 x'(t)^2 + y'(t)^2 &= (r'(t)\cos\theta(t) - r(t)\theta'(t)\sin\theta(t))^2 + (r'(t)\sin\theta(t) + r(t)\theta'(t)\cos\theta(t))^2 \\
 &= r'(t)^2(\cos^2\theta + \sin^2\theta) + r(t)^2\theta'(t)^2(\cos^2\theta + \sin^2\theta) \\
 &\quad - 2r'(t)r(t)\theta'(t)\cos\theta\sin\theta + 2r'(t)r(t)\theta'(t)\cos\theta\sin\theta \\
 &= r'(t)^2 + r(t)^2\theta'(t)^2, \text{ and so} \\
 L &= \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt \\
 &= \int_a^b \sqrt{r'(t)^2 + r(t)^2\theta'(t)^2 + z'(t)^2} dt
 \end{aligned}$$

QED

Example 1.44. Find the arc length L of the curve whose cylindrical coordinates are $r = e^t$, $\theta = t$ and $z = e^t$, for t over the interval $[0, 1]$.

Solution: Since $r'(t) = e^t$, $\theta'(t) = 1$ and $z'(t) = e^t$, then

$$\begin{aligned}
 L &= \int_0^1 \sqrt{r'(t)^2 + r(t)^2\theta'(t)^2 + z'(t)^2} dt \\
 &= \int_0^1 \sqrt{e^{2t} + e^{2t}(1) + e^{2t}} dt \\
 &= \int_0^1 e^t \sqrt{3} dt = \sqrt{3}(e - 1)
 \end{aligned}$$

Exercises

A

For Exercises 1-3, calculate the arc length of $\mathbf{f}(t)$ over the given interval.

1. $\mathbf{f}(t) = (3\cos 2t, 3\sin 2t, 3t)$ on $[0, \pi/2]$
2. $\mathbf{f}(t) = ((t^2 + 1)\cos t, (t^2 + 1)\sin t, 2\sqrt{2}t)$ on $[0, 1]$
3. $\mathbf{f}(t) = (2\cos 3t, 2\sin 3t, 2t^{3/2})$ on $[0, 1]$
4. Parametrize the curve from Exercise 1 by arc length.
5. Parametrize the curve from Exercise 3 by arc length.

B

6. Let $\mathbf{f}(t)$ be a differentiable curve such that $\mathbf{f}(t) \neq \mathbf{0}$ for all t . Show that

$$\frac{d}{dt} \left(\frac{\mathbf{f}(t)}{\|\mathbf{f}(t)\|} \right) = \frac{\mathbf{f}(t) \times (\mathbf{f}'(t) \times \mathbf{f}(t))}{\|\mathbf{f}(t)\|^3}.$$

Exercises 7-9 develop the moving frame field \mathbf{T} , \mathbf{N} , \mathbf{B} at a point on a curve.

7. Let $\mathbf{f}(t)$ be a smooth curve such that $\mathbf{f}'(t) \neq \mathbf{0}$ for all t . Then we can define the *unit tangent vector* \mathbf{T} by

$$\mathbf{T}(t) = \frac{\mathbf{f}'(t)}{\|\mathbf{f}'(t)\|}.$$

Show that

$$\mathbf{T}'(t) = \frac{\mathbf{f}'(t) \times (\mathbf{f}''(t) \times \mathbf{f}'(t))}{\|\mathbf{f}'(t)\|^3}.$$

8. Continuing Exercise 7, assume that $\mathbf{f}'(t)$ and $\mathbf{f}''(t)$ are not parallel. Then $\mathbf{T}'(t) \neq \mathbf{0}$ so we can define the *unit principal normal vector* \mathbf{N} by

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}.$$

Show that

$$\mathbf{N}(t) = \frac{\mathbf{f}'(t) \times (\mathbf{f}''(t) \times \mathbf{f}'(t))}{\|\mathbf{f}'(t)\| \|\mathbf{f}''(t) \times \mathbf{f}'(t)\|}.$$

9. Continuing Exercise 8, the *unit binormal vector* \mathbf{B} is defined by

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t).$$

Show that

$$\mathbf{B}(t) = \frac{\mathbf{f}'(t) \times \mathbf{f}''(t)}{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}.$$

Note: The vectors $\mathbf{T}(t)$, $\mathbf{N}(t)$ and $\mathbf{B}(t)$ form a right-handed system of mutually perpendicular unit vectors (called *orthonormal vectors*) at each point on the curve $\mathbf{f}(t)$.

10. Continuing Exercise 9, the *curvature* κ is defined by

$$\kappa(t) = \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{f}'(t)\|} = \frac{\|\mathbf{f}'(t) \times (\mathbf{f}''(t) \times \mathbf{f}'(t))\|}{\|\mathbf{f}'(t)\|^4}.$$

Show that

$$\kappa(t) = \frac{\|\mathbf{f}'(t) \times \mathbf{f}''(t)\|}{\|\mathbf{f}'(t)\|^3} \quad \text{and that} \quad \mathbf{T}'(t) = \|\mathbf{f}'(t)\| \kappa(t) \mathbf{N}(t).$$

Note: $\kappa(t)$ gives a sense of how “curved” the curve $\mathbf{f}(t)$ is at each point.

11. Find \mathbf{T} , \mathbf{N} , \mathbf{B} and κ at each point of the helix $\mathbf{f}(t) = (\cos t, \sin t, t)$.
12. Show that the arc length L of a curve whose spherical coordinates are $\rho = \rho(t)$, $\theta = \theta(t)$ and $\phi = \phi(t)$ for t in an interval $[a, b]$ is

$$L = \int_a^b \sqrt{\rho'(t)^2 + (\rho(t)^2 \sin^2 \phi(t)) \theta'(t)^2 + \rho(t)^2 \phi'(t)^2} dt.$$

4 Line and Surface Integrals

4.1 Line Integrals

In single-variable calculus you learned how to integrate a real-valued function $f(x)$ over an interval $[a, b]$ in \mathbb{R}^1 . This integral (usually called a *Riemann integral*) can be thought of as an integral over a *path* in \mathbb{R}^1 , since an interval (or collection of intervals) is really the only kind of “path” in \mathbb{R}^1 . You may also recall that if $f(x)$ represented the force applied along the x -axis to an object at position x in $[a, b]$, then the *work* W done in moving that object from position $x = a$ to $x = b$ was defined as the integral:

$$W = \int_a^b f(x) dx$$

In this section, we will see how to define the integral of a function (either real-valued or vector-valued) of two variables over a general path (i.e. a curve) in \mathbb{R}^2 . This definition will be motivated by the physical notion of work. We will begin with real-valued functions of two variables.

In physics, the intuitive idea of work is that

$$\text{Work} = \text{Force} \times \text{Distance} .$$

Suppose that we want to find the total amount W of work done in moving an object along a curve C in \mathbb{R}^2 with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$, with a force $f(x, y)$ which varies with the position (x, y) of the object and is applied in the direction of motion along C (see Figure 4.1.1 below).

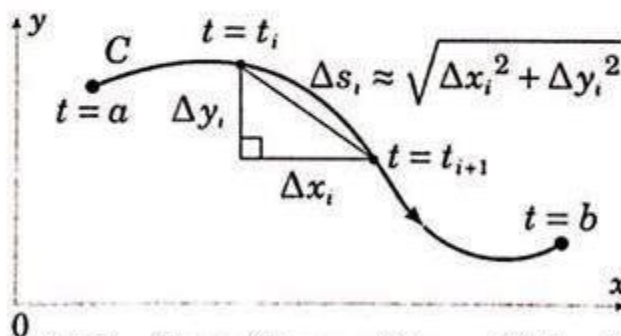


Figure 4.1.1 Curve $C : x = x(t), y = y(t)$ for t in $[a, b]$

We will assume for now that the function $f(x, y)$ is continuous and real-valued, so we only consider the magnitude of the force. Partition the interval $[a, b]$ as follows:

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b , \text{ for some integer } n \geq 2$$

As we can see from Figure 4.1.1, over a typical subinterval $[t_i, t_{i+1}]$ the distance Δs_i traveled along the curve is approximately $\sqrt{\Delta x_i^2 + \Delta y_i^2}$, by the Pythagorean Theorem. Thus, if the subinterval is small enough then the work done in moving the object along that piece of the curve is approximately

$$\text{Force} \times \text{Distance} \approx f(x_i, y_i) \sqrt{\Delta x_i^2 + \Delta y_i^2}, \quad (4.1)$$

where $(x_i, y_i) = (x(t_i^*), y(t_i^*))$ for some t_i^* in $[t_i, t_{i+1}]$, and so

$$W \approx \sum_{i=0}^{n-1} f(x_i, y_i) \sqrt{\Delta x_i^2 + \Delta y_i^2} \quad (4.2)$$

is approximately the total amount of work done over the entire curve. But since

$$\sqrt{\Delta x_i^2 + \Delta y_i^2} = \sqrt{\left(\frac{\Delta x_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta y_i}{\Delta t_i}\right)^2} \Delta t_i,$$

where $\Delta t_i = t_{i+1} - t_i$, then

$$W \approx \sum_{i=0}^{n-1} f(x_i, y_i) \sqrt{\left(\frac{\Delta x_i}{\Delta t_i}\right)^2 + \left(\frac{\Delta y_i}{\Delta t_i}\right)^2} \Delta t_i. \quad (4.3)$$

Taking the limit of that sum as the length of the largest subinterval goes to 0, the sum over all subintervals becomes the integral from $t = a$ to $t = b$, $\frac{\Delta x_i}{\Delta t_i}$ and $\frac{\Delta y_i}{\Delta t_i}$ become $x'(t)$ and $y'(t)$, respectively, and $f(x_i, y_i)$ becomes $f(x(t), y(t))$, so that

$$W = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2} dt. \quad (4.4)$$

The integral on the right side of the above equation gives us our idea of how to define, for any real-valued function $f(x, y)$, the integral of $f(x, y)$ along the curve C , called a *line integral*:

Definition 4.1. For a real-valued function $f(x, y)$ and a curve C in \mathbb{R}^2 , parametrized by $x = x(t)$, $y = y(t)$, $a \leq t \leq b$, the **line integral of $f(x, y)$ along C with respect to arc length** s is

$$\int_C f(x, y) ds = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2} dt. \quad (4.5)$$

The symbol ds is the differential of the arc length function

$$s = s(t) = \int_a^t \sqrt{x'(u)^2 + y'(u)^2} du, \quad (4.6)$$

which you may recognize from Section 1.9 as the length of the curve C over the interval $[a, t]$, for all t in $[a, b]$. That is,

$$ds = s'(t)dt = \sqrt{x'(t)^2 + y'(t)^2} dt, \quad (4.7)$$

by the Fundamental Theorem of Calculus.

For a general real-valued function $f(x, y)$, what does the line integral $\int_C f(x, y) ds$ represent? The preceding discussion of ds gives us a clue. You can think of differentials as infinitesimal lengths. So if you think of $f(x, y)$ as the height of a picket fence along C , then $f(x, y) ds$ can be thought of as approximately the area of a section of that fence over some infinitesimally small section of the curve, and thus the line integral $\int_C f(x, y) ds$ is the total area of that picket fence (see Figure 4.1.2).

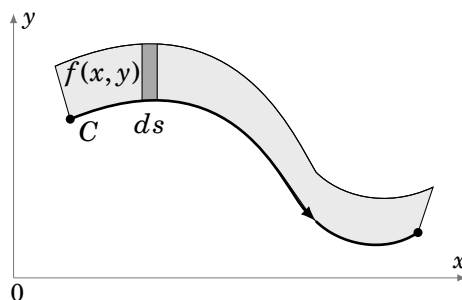


Figure 4.1.2 Area of shaded rectangle = height \times width $\approx f(x, y) ds$

Example 4.1. Use a line integral to show that the lateral surface area A of a right circular cylinder of radius r and height h is $2\pi rh$.

Solution: We will use the right circular cylinder with base circle C given by $x^2 + y^2 = r^2$ and with height h in the positive z direction (see Figure 4.1.3). Parametrize C as follows:

$$x = x(t) = r \cos t, \quad y = y(t) = r \sin t, \quad 0 \leq t \leq 2\pi$$

Let $f(x, y) = h$ for all (x, y) . Then

$$\begin{aligned} A &= \int_C f(x, y) ds = \int_a^b f(x(t), y(t)) \sqrt{x'(t)^2 + y'(t)^2} dt \\ &= \int_0^{2\pi} h \sqrt{(-r \sin t)^2 + (r \cos t)^2} dt \\ &= h \int_0^{2\pi} r \sqrt{\sin^2 t + \cos^2 t} dt \\ &= rh \int_0^{2\pi} 1 dt = 2\pi rh \end{aligned}$$

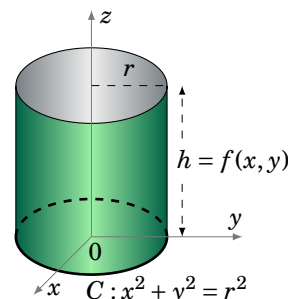


Figure 4.1.3

Note in Example 4.1 that if we had traversed the circle C twice, i.e. let t vary from 0 to 4π , then we would have gotten an area of $4\pi rh$, i.e. twice the desired area, even though the *curve* itself is still the same (namely, a circle of radius r). Also, notice that we traversed the circle in the counter-clockwise direction. If we had gone in the clockwise direction, using the parametrization

$$x = x(t) = r \cos(2\pi - t), \quad y = y(t) = r \sin(2\pi - t), \quad 0 \leq t \leq 2\pi, \quad (4.8)$$

then it is easy to verify (see Exercise 12) that the value of the line integral is unchanged.

In general, it can be shown (see Exercise 15) that reversing the direction in which a curve C is traversed leaves $\int_C f(x, y) ds$ unchanged, for any $f(x, y)$. If a curve C has a parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$, then denote by $-C$ the same curve as C but traversed in the opposite direction. Then $-C$ is parametrized by

$$x = x(a + b - t), \quad y = y(a + b - t), \quad a \leq t \leq b, \quad (4.9)$$

and we have

$$\int_C f(x, y) ds = \int_{-C} f(x, y) ds. \quad (4.10)$$

Notice that our definition of the line integral was with respect to the arc length parameter s . We can also define

$$\int_C f(x, y) dx = \int_a^b f(x(t), y(t)) x'(t) dt \quad (4.11)$$

as the *line integral of $f(x, y)$ along C with respect to x* , and

$$\int_C f(x, y) dy = \int_a^b f(x(t), y(t)) y'(t) dt \quad (4.12)$$

as the *line integral of $f(x, y)$ along C with respect to y* .

In the derivation of the formula for a line integral, we used the idea of work as force multiplied by distance. However, we know that force is actually a *vector*. So it would be helpful to develop a vector form for a line integral. For this, suppose that we have a function $\mathbf{f}(x, y)$ defined on \mathbb{R}^2 by

$$\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$$

for some continuous real-valued functions $P(x, y)$ and $Q(x, y)$ on \mathbb{R}^2 . Such a function \mathbf{f} is called a **vector field** on \mathbb{R}^2 . It is defined at *points* in \mathbb{R}^2 , and its values are *vectors* in \mathbb{R}^2 . For a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$, let

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$$

be the position vector for a point $(x(t), y(t))$ on C . Then $\mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$ and so

$$\begin{aligned}\int_C P(x, y) dx + \int_C Q(x, y) dy &= \int_a^b P(x(t), y(t)) x'(t) dt + \int_a^b Q(x(t), y(t)) y'(t) dt \\ &= \int_a^b (P(x(t), y(t)) x'(t) + Q(x(t), y(t)) y'(t)) dt \\ &= \int_a^b \mathbf{f}(x(t), y(t)) \cdot \mathbf{r}'(t) dt\end{aligned}$$

by definition of $\mathbf{f}(x, y)$. Notice that the function $\mathbf{f}(x(t), y(t)) \cdot \mathbf{r}'(t)$ is a *real-valued* function on $[a, b]$, so the last integral on the right looks somewhat similar to our earlier definition of a line integral. This leads us to the following definition:

Definition 4.2. For a vector field $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ and a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$, the **line integral of \mathbf{f} along C** is

$$\int_C \mathbf{f} \cdot d\mathbf{r} = \int_C P(x, y) dx + \int_C Q(x, y) dy \quad (4.13)$$

$$= \int_a^b \mathbf{f}(x(t), y(t)) \cdot \mathbf{r}'(t) dt, \quad (4.14)$$

where $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ is the position vector for points on C .

We use the notation $d\mathbf{r} = \mathbf{r}'(t)dt = dx\mathbf{i} + dy\mathbf{j}$ to denote the **differential** of the vector-valued function \mathbf{r} . The line integral in Definition 4.2 is often called a *line integral of a vector field* to distinguish it from the line integral in Definition 4.1 which is called a *line integral of a scalar field*. For convenience we will often write

$$\int_C P(x, y) dx + \int_C Q(x, y) dy = \int_C P(x, y) dx + Q(x, y) dy,$$

where it is understood that the line integral along C is being applied to both P and Q . The quantity $P(x, y)dx + Q(x, y)dy$ is known as a **differential form**. For a real-valued function $F(x, y)$, the **differential** of F is $dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$. A differential form $P(x, y)dx + Q(x, y)dy$ is called **exact** if it equals dF for some function $F(x, y)$.

Recall that if the points on a curve C have position vector $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, then $\mathbf{r}'(t)$ is a tangent vector to C at the point $(x(t), y(t))$ in the direction of increasing t (which we call the *direction of C*). Since C is a smooth curve, then $\mathbf{r}'(t) \neq \mathbf{0}$ on $[a, b]$ and hence

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$$

is the unit tangent vector to C at $(x(t), y(t))$. Putting Definitions 4.1 and 4.2 together we get the following theorem:

Theorem 4.1. For a vector field $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ and a curve C with a smooth parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ and position vector $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$,

$$\int_C \mathbf{f} \cdot d\mathbf{r} = \int_C \mathbf{f} \cdot \mathbf{T} ds, \quad (4.15)$$

where $\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$ is the unit tangent vector to C at $(x(t), y(t))$.

If the vector field $\mathbf{f}(x, y)$ represents the force moving an object along a curve C , then the work W done by this force is

$$W = \int_C \mathbf{f} \cdot \mathbf{T} ds = \int_C \mathbf{f} \cdot d\mathbf{r}. \quad (4.16)$$

Example 4.2. Evaluate $\int_C (x^2 + y^2) dx + 2xy dy$, where:

- (a) $C : x = t, \quad y = 2t, \quad 0 \leq t \leq 1$
 (b) $C : x = t, \quad y = 2t^2, \quad 0 \leq t \leq 1$

Solution: Figure 4.1.4 shows both curves.

(a) Since $x'(t) = 1$ and $y'(t) = 2$, then

$$\begin{aligned} \int_C (x^2 + y^2) dx + 2xy dy &= \int_0^1 ((x(t)^2 + y(t)^2)x'(t) + 2x(t)y(t)y'(t)) dt \\ &= \int_0^1 ((t^2 + 4t^2)(1) + 2t(2t)(2)) dt \\ &= \int_0^1 13t^2 dt \\ &= \left. \frac{13t^3}{3} \right|_0^1 = \frac{13}{3} \end{aligned}$$

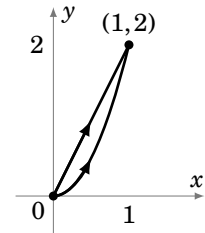


Figure 4.1.4

(b) Since $x'(t) = 1$ and $y'(t) = 4t$, then

$$\begin{aligned} \int_C (x^2 + y^2) dx + 2xy dy &= \int_0^1 ((x(t)^2 + y(t)^2)x'(t) + 2x(t)y(t)y'(t)) dt \\ &= \int_0^1 ((t^2 + 4t^4)(1) + 2t(2t^2)(4t)) dt \\ &= \int_0^1 (t^2 + 20t^4) dt \\ &= \left. \frac{t^3}{3} + 4t^5 \right|_0^1 = \frac{1}{3} + 4 = \frac{13}{3} \end{aligned}$$

So in both cases, if the vector field $\mathbf{f}(x, y) = (x^2 + y^2)\mathbf{i} + 2xy\mathbf{j}$ represents the force moving an object from $(0, 0)$ to $(1, 2)$ along the given curve C , then the work done is $\frac{13}{3}$. This may lead you to think that work (and more generally, the line integral of a vector field) is independent of the path taken. However, as we will see in the next section, this is not always the case.

Although we defined line integrals over a single smooth curve, if C is a *piecewise smooth curve*, that is

$$C = C_1 \cup C_2 \cup \dots \cup C_n$$

is the union of smooth curves C_1, \dots, C_n , then we can define

$$\int_C \mathbf{f} \cdot d\mathbf{r} = \int_{C_1} \mathbf{f} \cdot d\mathbf{r}_1 + \int_{C_2} \mathbf{f} \cdot d\mathbf{r}_2 + \dots + \int_{C_n} \mathbf{f} \cdot d\mathbf{r}_n$$

where each \mathbf{r}_i is the position vector of the curve C_i .

Example 4.3. Evaluate $\int_C (x^2 + y^2)dx + 2xydy$, where C is the polygonal path from $(0, 0)$ to $(0, 2)$ to $(1, 2)$.

Solution: Write $C = C_1 \cup C_2$, where C_1 is the curve given by $x = 0$, $y = t$, $0 \leq t \leq 2$ and C_2 is the curve given by $x = t$, $y = 2$, $0 \leq t \leq 1$ (see Figure 4.1.5). Then

$$\begin{aligned} \int_C (x^2 + y^2)dx + 2xydy &= \int_{C_1} (x^2 + y^2)dx + 2xydy \\ &\quad + \int_{C_2} (x^2 + y^2)dx + 2xydy \end{aligned}$$

$$= \int_0^2 ((0^2 + t^2)(0) + 2(0)t(1)) dt + \int_0^1 ((t^2 + 4)(1) + 2t(2)(0)) dt$$

$$= \int_0^2 0 dt + \int_0^1 (t^2 + 4) dt$$

$$= \frac{t^3}{3} + 4t \Big|_0^1 = \frac{1}{3} + 4 = \frac{13}{3}$$

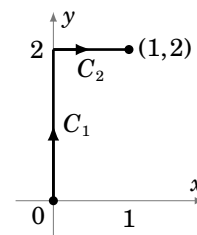


Figure 4.1.5

Line integral notation varies quite a bit. For example, in physics it is common to see the notation $\int_a^b \mathbf{f} \cdot d\mathbf{l}$, where it is understood that the limits of integration a and b are for the underlying parameter t of the curve, and the letter \mathbf{l} signifies length. Also, the formulation $\int_C \mathbf{f} \cdot \mathbf{T} ds$ from Theorem 4.1 is often preferred in physics since it emphasizes the idea of integrating the tangential component $\mathbf{f} \cdot \mathbf{T}$ of \mathbf{f} in the direction of \mathbf{T} (i.e. in the direction of C), which is a useful physical interpretation of line integrals.

Exercises

A

For Exercises 1-4, calculate $\int_C f(x, y) ds$ for the given function $f(x, y)$ and curve C .

1. $f(x, y) = xy$; $C : x = \cos t, y = \sin t, 0 \leq t \leq \pi/2$
2. $f(x, y) = \frac{x}{x^2 + 1}$; $C : x = t, y = 0, 0 \leq t \leq 1$
3. $f(x, y) = 2x + y$; C : polygonal path from $(0, 0)$ to $(3, 0)$ to $(3, 2)$
4. $f(x, y) = x + y^2$; C : path from $(2, 0)$ counterclockwise along the circle $x^2 + y^2 = 4$ to the point $(-2, 0)$ and then back to $(2, 0)$ along the x -axis
5. Use a line integral to find the lateral surface area of the part of the cylinder $x^2 + y^2 = 4$ below the plane $x + 2y + z = 6$ and above the xy -plane.

For Exercises 6-11, calculate $\int_C \mathbf{f} \cdot d\mathbf{r}$ for the given vector field $\mathbf{f}(x, y)$ and curve C .

6. $\mathbf{f}(x, y) = \mathbf{i} - \mathbf{j}$; $C : x = 3t, y = 2t, 0 \leq t \leq 1$
7. $\mathbf{f}(x, y) = y\mathbf{i} - x\mathbf{j}$; $C : x = \cos t, y = \sin t, 0 \leq t \leq 2\pi$
8. $\mathbf{f}(x, y) = x\mathbf{i} + y\mathbf{j}$; $C : x = \cos t, y = \sin t, 0 \leq t \leq 2\pi$
9. $\mathbf{f}(x, y) = (x^2 - y)\mathbf{i} + (x - y^2)\mathbf{j}$; $C : x = \cos t, y = \sin t, 0 \leq t \leq 2\pi$
10. $\mathbf{f}(x, y) = xy^2\mathbf{i} + xy^3\mathbf{j}$; C : the polygonal path from $(0, 0)$ to $(1, 0)$ to $(0, 1)$ to $(0, 0)$
11. $\mathbf{f}(x, y) = (x^2 + y^2)\mathbf{i}$; $C : x = 2 + \cos t, y = \sin t, 0 \leq t \leq 2\pi$

B

12. Verify that the value of the line integral in Example 4.1 is unchanged when using the parametrization of the circle C given in formulas (4.8).
13. Show that if $\mathbf{f} \perp \mathbf{r}'(t)$ at each point $\mathbf{r}(t)$ along a smooth curve C , then $\int_C \mathbf{f} \cdot d\mathbf{r} = 0$.
14. Show that if \mathbf{f} points in the same direction as $\mathbf{r}'(t)$ at each point $\mathbf{r}(t)$ along a smooth curve C , then $\int_C \mathbf{f} \cdot d\mathbf{r} = \int_C \|\mathbf{f}\| ds$.

C

15. Prove that $\int_C f(x, y) ds = \int_{-C} f(x, y) ds$. (Hint: Use formulas (4.9).)
16. Let C be a smooth curve with arc length L , and suppose that $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ is a vector field such that $\|\mathbf{f}(x, y)\| \leq M$ for all (x, y) on C . Show that $|\int_C \mathbf{f} \cdot d\mathbf{r}| \leq ML$. (Hint: Recall that $|\int_a^b g(x) dx| \leq \int_a^b |g(x)| dx$ for Riemann integrals.)
17. Prove that the Riemann integral $\int_a^b f(x) dx$ is a special case of a line integral.

4.2 Properties of Line Integrals

We know from the previous section that for line integrals of real-valued functions (scalar fields), reversing the direction in which the integral is taken along a curve does not change the value of the line integral:

$$\int_C f(x, y) ds = \int_{-C} f(x, y) ds \quad (4.17)$$

For line integrals of vector fields, however, the value does change. To see this, let $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ be a vector field, with P and Q continuously differentiable functions. Let C be a smooth curve parametrized by $x = x(t)$, $y = y(t)$, $a \leq t \leq b$, with position vector $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ (we will usually abbreviate this by saying that $C : \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ is a smooth curve). We know that the curve $-C$ traversed in the opposite direction is parametrized by $x = x(a + b - t)$, $y = y(a + b - t)$, $a \leq t \leq b$. Then

$$\begin{aligned} \int_{-C} P(x, y) dx &= \int_a^b P(x(a + b - t), y(a + b - t)) \frac{d}{dt}(x(a + b - t)) dt \\ &= \int_a^b P(x(a + b - t), y(a + b - t))(-x'(a + b - t)) dt \quad (\text{by the Chain Rule}) \\ &= \int_b^a P(x(u), y(u))(-x'(u))(-du) \quad (\text{by letting } u = a + b - t) \\ &= \int_b^a P(x(u), y(u))x'(u) du \\ &= -\int_a^b P(x(u), y(u))x'(u) du, \quad \text{since } \int_b^a = -\int_a^b, \text{ so} \\ \int_{-C} P(x, y) dx &= -\int_C P(x, y) dx \end{aligned}$$

since we are just using a different letter (u) for the line integral along C . A similar argument shows that

$$\int_{-C} Q(x, y) dy = -\int_C Q(x, y) dy,$$

and hence

$$\begin{aligned} \int_{-C} \mathbf{f} \cdot d\mathbf{r} &= \int_{-C} P(x, y) dx + \int_{-C} Q(x, y) dy \\ &= -\int_C P(x, y) dx - \int_C Q(x, y) dy \\ &= -\left(\int_C P(x, y) dx + \int_C Q(x, y) dy \right) \\ \int_{-C} \mathbf{f} \cdot d\mathbf{r} &= -\int_C \mathbf{f} \cdot d\mathbf{r}. \end{aligned} \quad (4.18)$$

The above formula can be interpreted in terms of the work done by a force $\mathbf{f}(x, y)$ (treated as a vector) moving an object along a curve C : the total work performed moving the object along C from its initial point to its terminal point, and then back to the initial point moving backwards along the same path, is zero. This is because when force is considered as a vector, direction is accounted for.

The preceding discussion shows the importance of always taking the *direction* of the curve into account when using line integrals of vector fields. For this reason, the curves in line integrals are sometimes referred to as *directed curves* or *oriented curves*.

Recall that our definition of a line integral required that we have a parametrization $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ for the curve C . But as we know, any curve has infinitely many parametrizations. So could we get a different value for a line integral using some other parametrization of C , say, $x = \tilde{x}(u)$, $y = \tilde{y}(u)$, $c \leq u \leq d$? If so, this would mean that our definition is not well-defined. Luckily, it turns out that the value of a line integral of a vector field is unchanged as long as the direction of the curve C is preserved by whatever parametrization is chosen:

Theorem 4.2. Let $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ be a vector field, and let C be a smooth curve parametrized by $x = x(t)$, $y = y(t)$, $a \leq t \leq b$. Suppose that $t = \alpha(u)$ for $c \leq u \leq d$, such that $a = \alpha(c)$, $b = \alpha(d)$, and $\alpha'(u) > 0$ on the open interval (c, d) (i.e. $\alpha(u)$ is strictly increasing on $[c, d]$). Then $\int_C \mathbf{f} \cdot d\mathbf{r}$ has the same value for the parametrizations $x = x(t)$, $y = y(t)$, $a \leq t \leq b$ and $x = \tilde{x}(u) = x(\alpha(u))$, $y = \tilde{y}(u) = y(\alpha(u))$, $c \leq u \leq d$.

Proof: Since $\alpha(u)$ is strictly increasing and maps $[c, d]$ onto $[a, b]$, then we know that $t = \alpha(u)$ has an inverse function $u = \alpha^{-1}(t)$ defined on $[a, b]$ such that $c = \alpha^{-1}(a)$, $d = \alpha^{-1}(b)$, and $\frac{du}{dt} = \frac{1}{\alpha'(u)}$. Also, $dt = \alpha'(u)du$, and by the Chain Rule

$$\tilde{x}'(u) = \frac{d\tilde{x}}{du} = \frac{d}{du}(x(\alpha(u))) = \frac{dx}{dt} \frac{dt}{du} = x'(t)\alpha'(u) \Rightarrow x'(t) = \frac{\tilde{x}'(u)}{\alpha'(u)}$$

so making the substitution $t = \alpha(u)$ gives

$$\begin{aligned} \int_a^b P(x(t), y(t))x'(t)dt &= \int_{\alpha^{-1}(a)}^{\alpha^{-1}(b)} P(x(\alpha(u)), y(\alpha(u))) \frac{\tilde{x}'(u)}{\alpha'(u)} (\alpha'(u)du) \\ &= \int_c^d P(\tilde{x}(u), \tilde{y}(u))\tilde{x}'(u)du, \end{aligned}$$

which shows that $\int_C P(x, y)dx$ has the same value for both parametrizations. A similar argument shows that $\int_C Q(x, y)dy$ has the same value for both parametrizations, and hence $\int_C \mathbf{f} \cdot d\mathbf{r}$ has the same value. **QED**

Notice that the condition $\alpha'(u) > 0$ in Theorem 4.2 means that the two parametrizations move along C in the same direction. That was *not* the case with the “reverse” parametrization for $-C$: for $u = a + b - t$ we have $t = \alpha(u) = a + b - u \Rightarrow \alpha'(u) = -1 < 0$.

Example 4.4. Evaluate the line integral $\int_C (x^2 + y^2)dx + 2xydy$ from Example 4.2, Section 4.1, along the curve $C : x = t, y = 2t^2, 0 \leq t \leq 1$, where $t = \sin u$ for $0 \leq u \leq \pi/2$.

Solution: First, we notice that $0 = \sin 0$, $1 = \sin(\pi/2)$, and $\frac{dt}{du} = \cos u > 0$ on $(0, \pi/2)$. So by Theorem 4.2 we know that if C is parametrized by

$$x = \sin u, \quad y = 2\sin^2 u, \quad 0 \leq u \leq \pi/2$$

then $\int_C (x^2 + y^2)dx + 2xydy$ should have the same value as we found in Example 4.2, namely $\frac{13}{3}$. And we can indeed verify this:

$$\begin{aligned} \int_C (x^2 + y^2)dx + 2xydy &= \int_0^{\pi/2} ((\sin^2 u + (2\sin^2 u)^2)\cos u + 2(\sin u)(2\sin^2 u)4\sin u \cos u) du \\ &= \int_0^{\pi/2} (\sin^2 u + 20\sin^4 u)\cos u du \\ &= \left. \frac{\sin^3 u}{3} + 4\sin^5 u \right|_0^{\pi/2} \\ &= \frac{1}{3} + 4 = \frac{13}{3} \end{aligned}$$

In other words, the line integral is unchanged whether t or u is the parameter for C .

By a **closed curve**, we mean a curve C whose initial point and terminal point are the same, i.e. for $C : x = x(t), y = y(t), a \leq t \leq b$, we have $(x(a), y(a)) = (x(b), y(b))$.

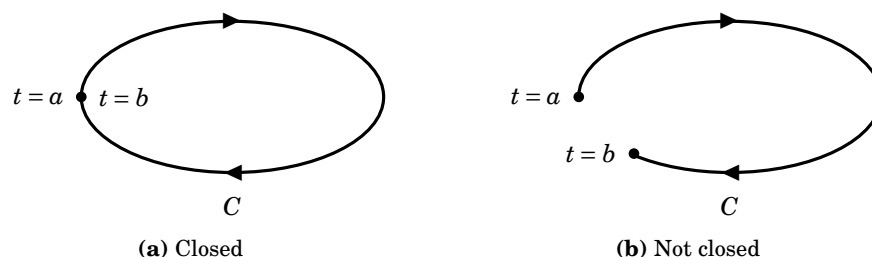


Figure 4.2.1 Closed vs nonclosed curves

A **simple closed curve** is a closed curve which does not intersect itself. Note that any closed curve can be regarded as a union of simple closed curves (think of the loops in a figure eight). We use the special notation

$$\oint_C f(x, y)ds \quad \text{and} \quad \oint_C \mathbf{f} \cdot d\mathbf{r}$$

to denote line integrals of scalar and vector fields, respectively, along closed curves. In some older texts you may see the notation \oint or \oint to indicate a line integral traversing a closed curve in a counterclockwise or clockwise direction, respectively.

So far, the examples we have seen of line integrals (e.g. Example 4.2) have had the same value for different curves joining the initial point to the terminal point. That is, the line integral has been independent of the path joining the two points. As we mentioned before, this is not always the case. The following theorem gives a necessary and sufficient condition for this *path independence*:

Theorem 4.3. In a region R , the line integral $\int_C \mathbf{f} \cdot d\mathbf{r}$ is independent of the path between any two points in R if and only if $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ for every closed curve C which is contained in R .

Proof: Suppose that $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ for every closed curve C which is contained in R . Let P_1 and P_2 be two distinct points in R . Let C_1 be a curve in R going from P_1 to P_2 , and let C_2 be another curve in R going from P_1 to P_2 , as in Figure 4.2.2.

Then $C = C_1 \cup -C_2$ is a closed curve in R (from P_1 to P_1), and so $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$. Thus,

$$\begin{aligned} 0 &= \oint_C \mathbf{f} \cdot d\mathbf{r} \\ &= \int_{C_1} \mathbf{f} \cdot d\mathbf{r} + \int_{-C_2} \mathbf{f} \cdot d\mathbf{r} \\ &= \int_{C_1} \mathbf{f} \cdot d\mathbf{r} - \int_{C_2} \mathbf{f} \cdot d\mathbf{r}, \text{ and so} \end{aligned}$$

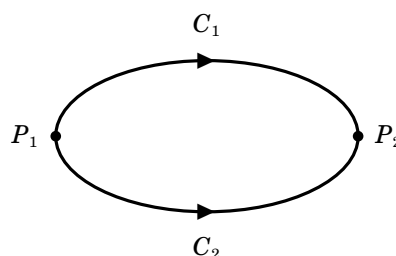


Figure 4.2.2

$\int_{C_1} \mathbf{f} \cdot d\mathbf{r} = \int_{C_2} \mathbf{f} \cdot d\mathbf{r}$. This proves path independence.

Conversely, suppose that the line integral $\int_C \mathbf{f} \cdot d\mathbf{r}$ is independent of the path between any two points in R . Let C be a closed curve contained in R . Let P_1 and P_2 be two distinct points on C . Let C_1 be a part of the curve C that goes from P_1 to P_2 , and let C_2 be the remaining part of C that goes from P_1 to P_2 , again as in Figure 4.2.2. Then by path independence we have

$$\begin{aligned} \int_{C_1} \mathbf{f} \cdot d\mathbf{r} &= \int_{C_2} \mathbf{f} \cdot d\mathbf{r} \\ \int_{C_1} \mathbf{f} \cdot d\mathbf{r} - \int_{C_2} \mathbf{f} \cdot d\mathbf{r} &= 0 \\ \int_{C_1} \mathbf{f} \cdot d\mathbf{r} + \int_{-C_2} \mathbf{f} \cdot d\mathbf{r} &= 0, \text{ so} \\ \oint_C \mathbf{f} \cdot d\mathbf{r} &= 0 \end{aligned}$$

since $C = C_1 \cup -C_2$.

QED

Clearly, the above theorem does not give a practical way to determine path independence, since it is impossible to check the line integrals around all possible closed curves in a region. What it mostly does is give an idea of the way in which line integrals behave, and how seemingly unrelated line integrals can be related (in this case, a specific line integral between two points and *all* line integrals around closed curves).

For a more practical method for determining path independence, we first need a version of the Chain Rule for multivariable functions:

Theorem 4.4. (Chain Rule) If $z = f(x, y)$ is a continuously differentiable function of x and y , and both $x = x(t)$ and $y = y(t)$ are differentiable functions of t , then z is a differentiable function of t , and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \quad (4.19)$$

at all points where the derivatives on the right are defined.

The proof is virtually identical to the proof of Theorem 2.2 from Section 2.4 (which uses the Mean Value Theorem), so we omit it.¹ We will now use this Chain Rule to prove the following *sufficient* condition for path independence of line integrals:

Theorem 4.5. Let $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ be a vector field in some region R , with P and Q continuously differentiable functions on R . Let C be a smooth curve in R parametrized by $x = x(t)$, $y = y(t)$, $a \leq t \leq b$. Suppose that there is a real-valued function $F(x, y)$ such that $\nabla F = \mathbf{f}$ on R . Then

$$\int_C \mathbf{f} \cdot d\mathbf{r} = F(B) - F(A), \quad (4.20)$$

where $A = (x(a), y(a))$ and $B = (x(b), y(b))$ are the endpoints of C . Thus, the line integral is independent of the path between its endpoints, since it depends only on the values of F at those endpoints.

Proof: By definition of $\int_C \mathbf{f} \cdot d\mathbf{r}$, we have

$$\begin{aligned} \int_C \mathbf{f} \cdot d\mathbf{r} &= \int_a^b (P(x(t), y(t))x'(t) + Q(x(t), y(t))y'(t)) dt \\ &= \int_a^b \left(\frac{\partial F}{\partial x} \frac{dx}{dt} + \frac{\partial F}{\partial y} \frac{dy}{dt} \right) dt \quad (\text{since } \nabla F = \mathbf{f} \Rightarrow \frac{\partial F}{\partial x} = P \text{ and } \frac{\partial F}{\partial y} = Q) \\ &= \int_a^b F'(x(t), y(t)) dt \quad (\text{by the Chain Rule in Theorem 4.4}) \\ &= F(x(t), y(t)) \Big|_a^b = F(B) - F(A) \end{aligned}$$

by the Fundamental Theorem of Calculus.

QED

¹See TAYLOR and MANN, § 6.5.

Theorem 4.5 can be thought of as the line integral version of the Fundamental Theorem of Calculus. A real-valued function $F(x, y)$ such that $\nabla F(x, y) = \mathbf{f}(x, y)$ is called a **potential** for \mathbf{f} . A **conservative** vector field is one which has a potential.

Example 4.5. Recall from Examples 4.2 and 4.3 in Section 4.1 that the line integral $\int_C (x^2 + y^2)dx + 2xydy$ was found to have the value $\frac{13}{3}$ for three different curves C going from the point $(0, 0)$ to the point $(1, 2)$. Use Theorem 4.5 to show that this line integral is indeed path independent.

Solution: We need to find a real-valued function $F(x, y)$ such that

$$\frac{\partial F}{\partial x} = x^2 + y^2 \quad \text{and} \quad \frac{\partial F}{\partial y} = 2xy.$$

Suppose that $\frac{\partial F}{\partial x} = x^2 + y^2$. Then we must have $F(x, y) = \frac{1}{3}x^3 + xy^2 + g(y)$ for some function $g(y)$. So $\frac{\partial F}{\partial y} = 2xy + g'(y)$ satisfies the condition $\frac{\partial F}{\partial y} = 2xy$ if $g'(y) = 0$, i.e. $g(y) = K$, where K is a constant. Since any choice for K will do (why?), we pick $K = 0$. Thus, a potential $F(x, y)$ for $\mathbf{f}(x, y) = (x^2 + y^2)\mathbf{i} + 2xy\mathbf{j}$ exists, namely

$$F(x, y) = \frac{1}{3}x^3 + xy^2.$$

Hence the line integral $\int_C (x^2 + y^2)dx + 2xydy$ is path independent.

Note that we can also verify that the value of the line integral of \mathbf{f} along any curve C going from $(0, 0)$ to $(1, 2)$ will always be $\frac{13}{3}$, since by Theorem 4.5

$$\int_C \mathbf{f} \cdot d\mathbf{r} = F(1, 2) - F(0, 0) = \frac{1}{3}(1)^3 + (1)(2)^2 - (0 + 0) = \frac{1}{3} + 4 = \frac{13}{3}.$$

A consequence of Theorem 4.5 in the special case where C is a closed curve, so that the endpoints A and B are the same point, is the following important corollary:

Corollary 4.6. If a vector field \mathbf{f} has a potential in a region R , then $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ for any closed curve C in R (i.e. $\oint_C \nabla F \cdot d\mathbf{r} = 0$ for any real-valued function $F(x, y)$).

Example 4.6. Evaluate $\oint_C xdx + ydy$ for $C : x = 2\cos t, y = 3\sin t, 0 \leq t \leq 2\pi$.

Solution: The vector field $\mathbf{f}(x, y) = x\mathbf{i} + y\mathbf{j}$ has a potential $F(x, y)$:

$$\begin{aligned} \frac{\partial F}{\partial x} = x &\Rightarrow F(x, y) = \frac{1}{2}x^2 + g(y), \text{ so} \\ \frac{\partial F}{\partial y} = y &\Rightarrow g'(y) = y \Rightarrow g(y) = \frac{1}{2}y^2 + K \end{aligned}$$

for any constant K , so $F(x, y) = \frac{1}{2}x^2 + \frac{1}{2}y^2$ is a potential for $\mathbf{f}(x, y)$. Thus,

$$\oint_C x dx + y dy = \oint_C \mathbf{f} \cdot d\mathbf{r} = 0$$

by Corollary 4.6, since the curve C is closed (it is the ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$).

Exercises

A

1. Evaluate $\oint_C (x^2 + y^2) dx + 2xy dy$ for $C : x = \cos t, y = \sin t, 0 \leq t \leq 2\pi$.
2. Evaluate $\int_C (x^2 + y^2) dx + 2xy dy$ for $C : x = \cos t, y = \sin t, 0 \leq t \leq \pi$.
3. Is there a potential $F(x, y)$ for $\mathbf{f}(x, y) = y\mathbf{i} - x\mathbf{j}$? If so, find one.
4. Is there a potential $F(x, y)$ for $\mathbf{f}(x, y) = x\mathbf{i} - y\mathbf{j}$? If so, find one.
5. Is there a potential $F(x, y)$ for $\mathbf{f}(x, y) = xy^2\mathbf{i} + x^3y\mathbf{j}$? If so, find one.

B

6. Let $\mathbf{f}(x, y)$ and $\mathbf{g}(x, y)$ be vector fields, let a and b be constants, and let C be a curve in \mathbb{R}^2 . Show that

$$\int_C (a\mathbf{f} \pm b\mathbf{g}) \cdot d\mathbf{r} = a \int_C \mathbf{f} \cdot d\mathbf{r} \pm b \int_C \mathbf{g} \cdot d\mathbf{r}.$$

7. Let C be a curve whose arc length is L . Show that $\int_C 1 ds = L$.
8. Let $f(x, y)$ and $g(x, y)$ be continuously differentiable real-valued functions in a region R . Show that

$$\oint_C f \nabla g \cdot d\mathbf{r} = - \oint_C g \nabla f \cdot d\mathbf{r}$$

for any closed curve C in R . (Hint: Use Exercise 21 in Section 2.4.)

9. Let $\mathbf{f}(x, y) = \frac{-y}{x^2+y^2}\mathbf{i} + \frac{x}{x^2+y^2}\mathbf{j}$ for all $(x, y) \neq (0, 0)$, and $C : x = \cos t, y = \sin t, 0 \leq t \leq 2\pi$.
 - (a) Show that $\mathbf{f} = \nabla F$, for $F(x, y) = \tan^{-1}(y/x)$.
 - (b) Show that $\oint_C \mathbf{f} \cdot d\mathbf{r} = 2\pi$. Does this contradict Corollary 4.6? Explain.

C

10. Let $g(x)$ and $h(y)$ be differentiable functions, and let $\mathbf{f}(x, y) = h(y)\mathbf{i} + g(x)\mathbf{j}$. Can \mathbf{f} have a potential $F(x, y)$? If so, find it. You may assume that F would be smooth. (Hint: Consider the mixed partial derivatives of F .)

4.3 Green's Theorem

We will now see a way of evaluating the line integral of a *smooth* vector field around a simple closed curve. A vector field $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ is **smooth** if its component functions $P(x, y)$ and $Q(x, y)$ are smooth. We will use *Green's Theorem* (sometimes called *Green's Theorem in the plane*) to relate the *line* integral around a closed curve with a *double* integral over the region inside the curve:

Theorem 4.7. (Green's Theorem) Let R be a region in \mathbb{R}^2 whose boundary is a simple closed curve C which is piecewise smooth. Let $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ be a smooth vector field defined on both R and C . Then

$$\oint_C \mathbf{f} \cdot d\mathbf{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA, \quad (4.21)$$

where C is traversed so that R is always on the left side of C .

Proof: We will prove the theorem in the case for a *simple* region R , that is, where the boundary curve C can be written as $C = C_1 \cup C_2$ in two distinct ways:

$$C_1 = \text{the curve } y = y_1(x) \text{ from the point } X_1 \text{ to the point } X_2 \quad (4.22)$$

$$C_2 = \text{the curve } y = y_2(x) \text{ from the point } X_2 \text{ to the point } X_1, \quad (4.23)$$

where X_1 and X_2 are the points on C farthest to the left and right, respectively; and

$$C_1 = \text{the curve } x = x_1(y) \text{ from the point } Y_2 \text{ to the point } Y_1 \quad (4.24)$$

$$C_2 = \text{the curve } x = x_2(y) \text{ from the point } Y_1 \text{ to the point } Y_2, \quad (4.25)$$

where Y_1 and Y_2 are the lowest and highest points, respectively, on C . See Figure 4.3.1.

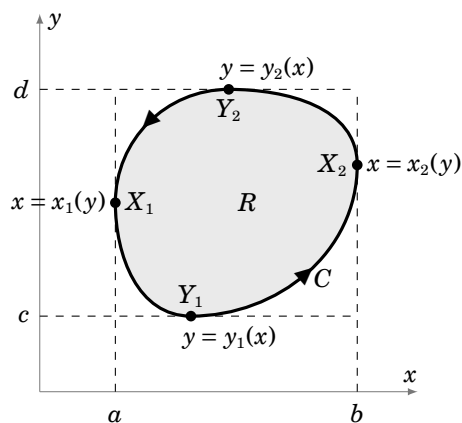


Figure 4.3.1

Integrate $P(x, y)$ around C using the representation $C = C_1 \cup C_2$ given by (4.23) and (4.24).

Since $y = y_1(x)$ along C_1 (as x goes from a to b) and $y = y_2(x)$ along C_2 (as x goes from b to a), as we see from Figure 4.3.1, then we have

$$\begin{aligned}
 \oint_C P(x, y) dx &= \int_{C_1} P(x, y) dx + \int_{C_2} P(x, y) dx \\
 &= \int_a^b P(x, y_1(x)) dx + \int_b^a P(x, y_2(x)) dx \\
 &= \int_a^b P(x, y_1(x)) dx - \int_a^b P(x, y_2(x)) dx \\
 &= - \int_a^b (P(x, y_2(x)) - P(x, y_1(x))) dx \\
 &= - \int_a^b \left(P(x, y) \Big|_{y=y_1(x)}^{y=y_2(x)} \right) dx \\
 &= - \int_a^b \int_{y_1(x)}^{y_2(x)} \frac{\partial P(x, y)}{\partial y} dy dx \quad (\text{by the Fundamental Theorem of Calculus}) \\
 &= - \iint_R \frac{\partial P}{\partial y} dA .
 \end{aligned}$$

Likewise, integrate $Q(x, y)$ around C using the representation $C = C_1 \cup C_2$ given by (4.25) and (4.26). Since $x = x_1(y)$ along C_1 (as y goes from d to c) and $x = x_2(y)$ along C_2 (as y goes from c to d), as we see from Figure 4.3.1, then we have

$$\begin{aligned}
 \oint_C Q(x, y) dy &= \int_{C_1} Q(x, y) dy + \int_{C_2} Q(x, y) dy \\
 &= \int_d^c Q(x_1(y), y) dy + \int_c^d Q(x_2(y), y) dy \\
 &= - \int_c^d Q(x_1(y), y) dy + \int_c^d Q(x_2(y), y) dy \\
 &= \int_c^d (Q(x_2(y), y) - Q(x_1(y), y)) dy \\
 &= \int_c^d \left(Q(x, y) \Big|_{x=x_1(y)}^{x=x_2(y)} \right) dy \\
 &= \int_c^d \int_{x_1(y)}^{x_2(y)} \frac{\partial Q(x, y)}{\partial x} dx dy \quad (\text{by the Fundamental Theorem of Calculus}) \\
 &= \iint_R \frac{\partial Q}{\partial x} dA , \text{ and so}
 \end{aligned}$$

$$\begin{aligned}
\oint_C \mathbf{f} \cdot d\mathbf{r} &= \oint_C P(x, y) dx + \oint_C Q(x, y) dy \\
&= - \iint_R \frac{\partial P}{\partial y} dA + \iint_R \frac{\partial Q}{\partial x} dA \\
&= \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA .
\end{aligned}$$

QED

Though we proved Green's Theorem only for a simple region R , the theorem can also be proved for more general regions (say, a union of simple regions).²

Example 4.7. Evaluate $\oint_C (x^2 + y^2) dx + 2xy dy$, where C is the boundary (traversed counter-clockwise) of the region $R = \{(x, y) : 0 \leq x \leq 1, 2x^2 \leq y \leq 2x\}$.

Solution: R is the shaded region in Figure 4.3.2. By Green's Theorem, for $P(x, y) = x^2 + y^2$ and $Q(x, y) = 2xy$, we have

$$\begin{aligned}
\oint_C (x^2 + y^2) dx + 2xy dy &= \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA \\
&= \iint_R (2y - 2y) dA = \iint_R 0 dA = 0 .
\end{aligned}$$

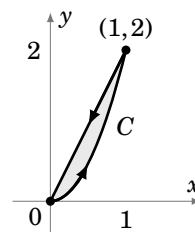


Figure 4.3.2

We actually already knew that the answer was zero. Recall from Example 4.5 in Section 4.2 that the vector field $\mathbf{f}(x, y) = (x^2 + y^2)\mathbf{i} + 2xy\mathbf{j}$ has a potential function $F(x, y) = \frac{1}{3}x^3 + xy^2$, and so $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ by Corollary 4.6.

Example 4.8. Let $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$, where

$$P(x, y) = \frac{-y}{x^2 + y^2} \quad \text{and} \quad Q(x, y) = \frac{x}{x^2 + y^2} ,$$

and let $R = \{(x, y) : 0 < x^2 + y^2 \leq 1\}$. For the boundary curve $C : x^2 + y^2 = 1$, traversed counter-clockwise, it was shown in Exercise 9(b) in Section 4.2 that $\oint_C \mathbf{f} \cdot d\mathbf{r} = 2\pi$. But

$$\frac{\partial Q}{\partial x} = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial P}{\partial y} \Rightarrow \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_R 0 dA = 0 .$$

This would seem to contradict Green's Theorem. However, note that R is not the *entire* region enclosed by C , since the point $(0, 0)$ is not contained in R . That is, R has a "hole" at the origin, so Green's Theorem does not apply.

²See TAYLOR and MANN, § 15.31 for a discussion of some of the difficulties involved when the boundary curve is "complicated".

If we modify the region R to be the *annulus* $R = \{(x, y) : 1/4 \leq x^2 + y^2 \leq 1\}$ (see Figure 4.3.3), and take the “boundary” C of R to be $C = C_1 \cup C_2$, where C_1 is the unit circle $x^2 + y^2 = 1$ traversed counterclockwise and C_2 is the circle $x^2 + y^2 = 1/4$ traversed *clockwise*, then it can be shown (see Exercise 8) that

$$\oint_C \mathbf{f} \cdot d\mathbf{r} = 0.$$

We would still have $\iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = 0$, so for this R we would have

$$\oint_C \mathbf{f} \cdot d\mathbf{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA,$$

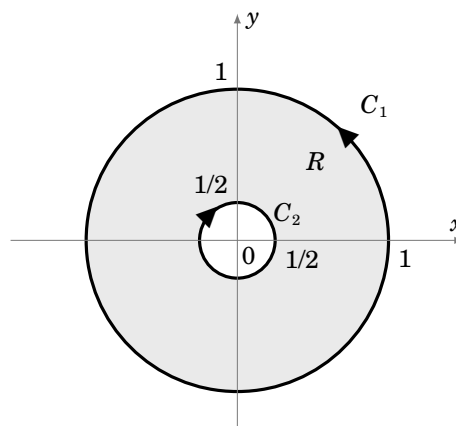


Figure 4.3.3 The annulus R

which shows that Green’s Theorem holds for the annular region R .

It turns out that Green’s Theorem can be extended to *multiply connected* regions, that is, regions like the annulus in Example 4.8, which have one or more *regions* cut out from the interior, as opposed to discrete *points* being cut out. For such regions, the “outer” boundary and the “inner” boundaries are traversed so that R is always on the left side.

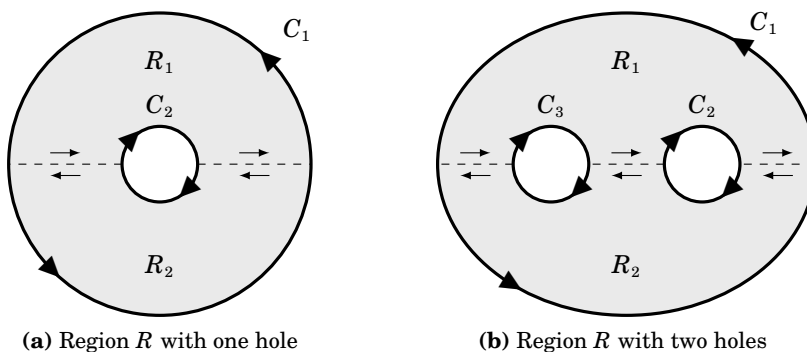


Figure 4.3.4 Multiply connected regions

The intuitive idea for why Green’s Theorem holds for multiply connected regions is shown in Figure 4.3.4 above. The idea is to cut “slits” between the boundaries of a multiply connected region R so that R is divided into subregions which do *not* have any “holes”. For example, in Figure 4.3.4(a) the region R is the union of the regions R_1 and R_2 , which are divided by the slits indicated by the dashed lines. Those slits are part of the boundary of both R_1 and R_2 , and we traverse then in the manner indicated by the arrows. Notice that along each slit the boundary of R_1 is traversed in the opposite direction as that of R_2 , which

means that the line integrals of \mathbf{f} along those slits cancel each other out. Since R_1 and R_2 do not have holes in them, then Green's Theorem holds in each subregion, so that

$$\oint_{\text{bdy of } R_1} \mathbf{f} \cdot d\mathbf{r} = \iint_{R_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA \quad \text{and} \quad \oint_{\text{bdy of } R_2} \mathbf{f} \cdot d\mathbf{r} = \iint_{R_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA .$$

But since the line integrals along the slits cancel out, we have

$$\oint_{C_1 \cup C_2} \mathbf{f} \cdot d\mathbf{r} = \oint_{\text{bdy of } R_1} \mathbf{f} \cdot d\mathbf{r} + \oint_{\text{bdy of } R_2} \mathbf{f} \cdot d\mathbf{r} ,$$

and so

$$\oint_{C_1 \cup C_2} \mathbf{f} \cdot d\mathbf{r} = \iint_{R_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA + \iint_{R_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA ,$$

which shows that Green's Theorem holds in the region R . A similar argument shows that the theorem holds in the region with two holes shown in Figure 4.3.4(b).

We know from Corollary 4.6 that when a smooth vector field $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ on a region R (whose boundary is a piecewise smooth, simple closed curve C) has a potential in R , then $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$. And if the potential $F(x, y)$ is smooth in R , then $\frac{\partial F}{\partial x} = P$ and $\frac{\partial F}{\partial y} = Q$, and so we know that

$$\frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y} \Rightarrow \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad \text{in } R .$$

Conversely, if $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ in R then

$$\oint_C \mathbf{f} \cdot d\mathbf{r} = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \iint_R 0 dA = 0 .$$

For a **simply connected** region R (i.e. a region with no holes), the following can be shown:

The following statements are equivalent for a simply connected region R in \mathbb{R}^2 :

- (a) $\mathbf{f}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ has a smooth potential $F(x, y)$ in R
- (b) $\int_C \mathbf{f} \cdot d\mathbf{r}$ is independent of the path for any curve C in R
- (c) $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$ for every simple closed curve C in R
- (d) $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ in R (in this case, the differential form $P dx + Q dy$ is exact)

Exercises

A

For Exercises 1-4, use Green's Theorem to evaluate the given line integral around the curve C , traversed counterclockwise.

1. $\oint_C (x^2 - y^2)dx + 2xydy$; C is the boundary of $R = \{(x, y) : 0 \leq x \leq 1, 2x^2 \leq y \leq 2x\}$
2. $\oint_C x^2ydx + 2xydy$; C is the boundary of $R = \{(x, y) : 0 \leq x \leq 1, x^2 \leq y \leq x\}$
3. $\oint_C 2ydx - 3xdy$; C is the circle $x^2 + y^2 = 1$
4. $\oint_C (e^{x^2} + y^2)dx + (e^{y^2} + x^2)dy$; C is the boundary of the triangle with vertices $(0,0)$, $(4,0)$ and $(0,4)$
5. Is there a potential $F(x, y)$ for $\mathbf{f}(x, y) = (y^2 + 3x^2)\mathbf{i} + 2xy\mathbf{j}$? If so, find one.
6. Is there a potential $F(x, y)$ for $\mathbf{f}(x, y) = (x^3 \cos(xy) + 2x \sin(xy))\mathbf{i} + x^2y \cos(xy)\mathbf{j}$? If so, find one.
7. Is there a potential $F(x, y)$ for $\mathbf{f}(x, y) = (8xy + 3)\mathbf{i} + 4(x^2 + y)\mathbf{j}$? If so, find one.
8. Show that for any constants a, b and any closed simple curve C , $\oint_C a dx + b dy = 0$.

B

9. For the vector field \mathbf{f} as in Example 4.8, show directly that $\oint_C \mathbf{f} \cdot d\mathbf{r} = 0$, where C is the boundary of the annulus $R = \{(x, y) : 1/4 \leq x^2 + y^2 \leq 1\}$ traversed so that R is always on the left.
10. Evaluate $\oint_C e^x \sin y dx + (y^3 + e^x \cos y)dy$, where C is the boundary of the rectangle with vertices $(1, -1)$, $(1, 1)$, $(-1, 1)$ and $(-1, -1)$, traversed counterclockwise.

C

11. For a region R bounded by a simple closed curve C , show that the area A of R is

$$A = -\oint_C y dx = \oint_C x dy = \frac{1}{2} \oint_C x dy - y dx,$$

where C is traversed so that R is always on the left. (*Hint: Use Green's Theorem and the fact that $A = \iint_R 1 dA$.*)

Chapter eight

Vector Algebra and functions

Overview

Many of you will know a good deal already about **Vector Algebra** — how to add and subtract vectors, how to take scalar and vector products of vectors, and something of how to describe geometric and physical entities using vectors. This course will remind you about that good stuff, but goes on to introduce you to the subject of **Vector Calculus** which, like it says on the can, combines vector algebra with calculus.

To give you a feeling for the issues, suppose you were interested in the temperature T of water in a river. Temperature T is a *scalar*, and will certainly be a function of a position vector $\mathbf{x} = (x, y, z)$ and may also be a function of time t : $T = T(\mathbf{x}, t)$. It is a scalar field.

Suppose now that you kept y, z, t constant, and asked what is the change in temperature as you move a small amount in x ? No doubt you'd be interested in calculating $\partial T / \partial x$. Similarly if you kept the point fixed, and asked how does the temperature change of time, you would be interested in $\partial T / \partial t$.

But why restrict ourselves to movements up-down, left-right, etc? Suppose you wanted to know what the change in temperature along an arbitrary direction. You would be interested in

$$\frac{\partial T}{\partial \mathbf{x}}$$

but how would you calculate that? Is $\partial T / \partial \mathbf{x}$ a vector or a scalar?

Now let's dive into the flow. At each point \mathbf{x} in the stream, at each time t , there will be a stream velocity $\mathbf{v}(\mathbf{x}, t)$. The local stream velocity can be viewed directly using modern techniques such as laser Doppler anemometry, or traditional techniques such as throwing twigs in. The point now is that \mathbf{v} is a function that has the same four input variables as temperature did, but its output result is a vector. We may be interested in places \mathbf{x} where the stream suddenly accelerates, or vortices where the stream curls around dangerously. That is, we will be interested in finding the acceleration of the stream, the gradient of its velocity. We may be interested in the magnitude of the acceleration (a scalar). Equally, we may be interested in the acceleration as a vector, so that we can apply Newton's law and figure out the force.

This is the stuff of vector calculus.

Grey book

Vector algebra: scalar and vector products; scalar and vector triple products; geometric applications. Differentiation of a vector function; scalar and vector fields. Gradient, divergence and curl – definitions and physical interpretations; product formulae; curvilinear coordinates. Gauss' and Stokes' theorems and evaluation of integrals over lines, surfaces and volumes. Derivation of continuity equations and Laplace's equation in Cartesian, cylindrical and spherical coordinate systems.

Course Content

- Introduction and revision of elementary concepts, scalar product, vector product.
- Triple products, multiple products, applications to geometry.
- Differentiation and integration of vector functions of a single variable.
- Curvilinear coordinate systems. Line, surface and volume integrals.
- Vector operators.
- Vector Identities.
- Gauss' and Stokes' Theorems.
- Engineering Applications.

Learning Outcomes

You should be comfortable with expressing systems (especially those in 2 and 3 dimensions) using vector quantities and manipulating these vectors without necessarily going back to some underlying coordinates.

You should have a sound grasp of the concept of a vector field, and be able to link this idea to descriptions of various physical phenomena.

You should have a good intuition of the physical meaning of the various vector calculus operators and the important related theorems. You should be able to interpret the formulae describing physical systems in terms of this intuition.

References

Although these notes cover the material you need to know you should, wider reading is essential. Different explanations and different diagrams in books will give you the perspective to glue everything together, and further worked examples give you the confidence to tackle the tute sheets.

- J Heading, "Mathematical Methods in Science and Engineering", 2nd ed., Ch.13, (Arnold).
- G Stephenson, "Mathematical Methods for Science Students", 2nd ed., Ch.19, (Longman).
- E Kreyszig, "Advanced Engineering Mathematics", 6th ed., Ch.6, (Wiley).
- K F Riley, M. P. Hobson and S. J. Bence, "Mathematical Methods for the Physics and Engineering" Chs.6, 8 and 9, (CUP).
- A J M Spencer, et. al. "Engineering Mathematics", Vol.1, Ch.6, (Van Nostrand Reinhold).
- H M Schey, "Div, Grad, Curl and all that", Norton

Lecture 1

Vector Algebra

1.1 Vectors

Many physical quantities, such as mass, time, temperature are fully specified by one number or magnitude. They are **scalars**. But other quantities require more than one number to describe them. They are **vectors**. You have already met vectors in their more pure mathematical sense in your course on linear algebra (matrices and vectors), but often in the physical world, these numbers specify a **magnitude** and a **direction** — a total of two numbers in a 2D planar world, and three numbers in 3D.

Obvious examples are velocity, acceleration, electric field, and force. Below, probably all our examples will be of these “magnitude and direction” vectors, but we should not forget that many of the results extend to the wider realm of vectors.

There are three slightly different types of vectors:

- **Free vectors:** In many situations *only* the magnitude and direction of a vector are important, and we can *translate* them at will (with 3 degrees of freedom for a vector in 3-dimensions).
- **Sliding vectors:** In mechanics the line of action of a force is often important for deriving moments. The force vector can slide with 1 degree of freedom.
- **Bound or position vectors:** When describing lines, curves etc in space, it is obviously important that the origin and head of the vector are not translated about arbitrarily. The origins of position vectors all coincide at an overall origin O .

One of the advantages of using vectors is that it frees much of the analysis from the restriction of arbitrarily imposed coordinate frames. For example, if two free vectors are equal we need only say that their magnitudes and directions are equal, and that can be done with a drawing that is independent of any coordinate system.

However, coordinate systems are ultimately useful, so it is useful to introduce the idea of vector components. Try to spot things in the notes that are independent

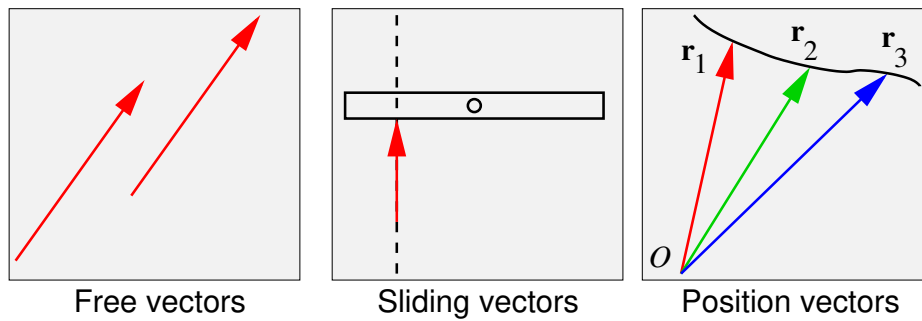


Figure 1.1:

of coordinate system.

1.1.1 Vector elements or components in a coordinate frame

A method of representing a vector is to list the values of its elements or components in a sufficient number of different (preferably mutually perpendicular) directions, depending on the dimension of the vector. These specified directions define a **coordinate frame**. In this course we will mostly restrict our attention to the 3-dimensional Cartesian coordinate frame $O(x, y, z)$. When we come to examine vector fields later in the course you will use curvilinear coordinate frames, especially 3D spherical and cylindrical polars, and 2D plane polar, coordinate systems.

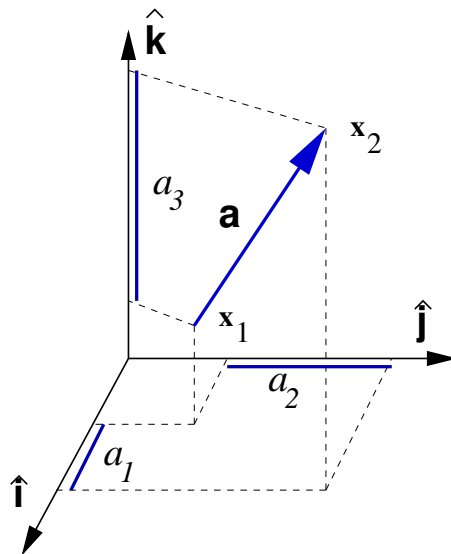


Figure 1.2: Vector components.

In a Cartesian coordinate frame we write

$$\mathbf{a} = [a_1, a_2, a_3] = [x_2 - x_1, y_2 - y_1, z_2 - z_1] \quad \text{or} \quad \mathbf{a} = [a_x, a_y, a_z]$$

as sketched in Figure 1.2. Defining $\hat{i}, \hat{j}, \hat{k}$ as unit vectors in the x, y, z directions

$$\hat{i} = [1, 0, 0] \quad \hat{j} = [0, 1, 0] \quad \hat{k} = [0, 0, 1]$$

1.1. VECTORS

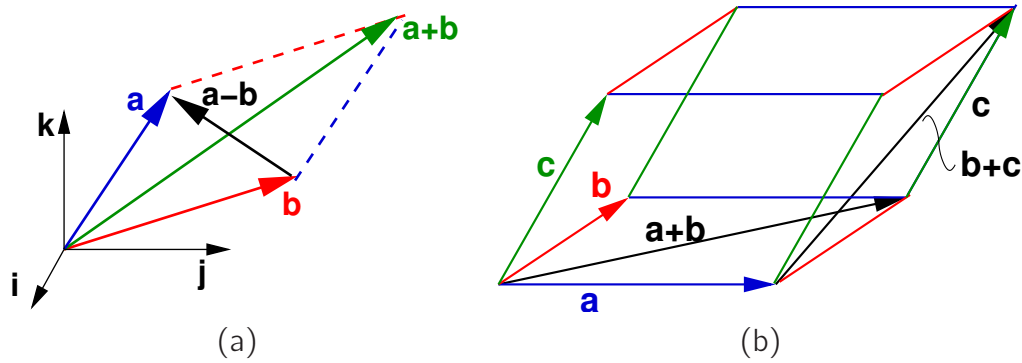


Figure 1.3: (a) Addition of two vectors is commutative, but subtraction isn't. Note that the coordinate frame is irrelevant. (b) Addition of three vectors is associative.

we could also write

$$\mathbf{a} = a_1 \hat{\mathbf{i}} + a_2 \hat{\mathbf{j}} + a_3 \hat{\mathbf{k}} \quad .$$

Although we will be most often dealing with vectors in 3-space, you should not think that general vectors are limited to three components.

In these notes we will use bold font to represent vectors **a**, **w**. In your written work, underline the vector symbol a, w and be **meticulous** about doing so. We shall use the hat to denote a unit vector.

1.1.2 Vector equality

Two free vectors are said to be equal iff their lengths and directions are the same. If we use a coordinate frame, we might say that corresponding components of the two vectors must be equal. This definition of equality will also do for position vectors, but for sliding vectors we must add that the line of action must be identical too.

1.1.3 Vector magnitude and unit vectors

Provided we use an orthogonal coordinate system, the magnitude of a 3-vector is

$$a = |\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

To find the unit vector in the direction of **a**, simply divide by its magnitude

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} \quad .$$

1.1.4 Vector Addition and subtraction

Vectors are added/subtracted by adding/subtracting corresponding components, exactly as for matrices. Thus

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1, \quad a_2 + b_2, \quad a_3 + b_3]$$

Addition follows the parallelogram construction of Figure 1.3(a). Subtraction ($\mathbf{a} - \mathbf{b}$) is defined as the addition ($\mathbf{a} + (-\mathbf{b})$). It is useful to remember that the vector $\mathbf{a} - \mathbf{b}$ goes from \mathbf{b} to \mathbf{a} .

The following results follow immediately from the above definition of vector addition:

- (a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (commutativity) (Figure 1.3(a))
- (b) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) = \mathbf{a} + \mathbf{b} + \mathbf{c}$ (associativity) (Figure 1.3(b))
- (c) $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$, where the zero vector is $\mathbf{0} = [0, 0, 0]$.
- (d) $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$

1.1.5 Multiplication of a vector by a scalar. (NOT the scalar product!)

Just as for matrices, multiplication of a vector \mathbf{a} by a scalar c is defined as multiplication of each component by c , so that

$$c\mathbf{a} = [ca_1, ca_2, ca_3].$$

It follows that:

$$|c\mathbf{a}| = \sqrt{(ca_1)^2 + (ca_2)^2 + (ca_3)^2} = |c||\mathbf{a}|.$$

The direction of the vector will reverse if c is negative, but otherwise is unaffected. (By the way, a vector where the sign is uncertain is called a director.)

♣ Example

Q. Coulomb's law states that the electrostatic force on charged particle Q due to another charged particle q_1 is

$$\mathbf{F} = K \frac{Qq_1}{r^2} \hat{\mathbf{e}}_r$$

where \mathbf{r} is the vector from q_1 to Q and $\hat{\mathbf{r}}$ is the unit vector in that same direction. (Note that the rule "unlike charges attract, like charges repel" is built into this formula.) The force between two particles is not modified by the presence of other charged particles.

Hence write down an expression for the force on Q at \mathbf{R} due to N charges q_i at \mathbf{r}_i .

A. The vector from q_i to Q is $\mathbf{R} - \mathbf{r}_i$. The unit vector in that direction is $(\mathbf{R} - \mathbf{r}_i)/|\mathbf{R} - \mathbf{r}_i|$, so the resultant force is

$$\mathbf{F}(\mathbf{R}) = \sum_{i=1}^N K \frac{Qq_i}{|\mathbf{R} - \mathbf{r}_i|^3} (\mathbf{R} - \mathbf{r}_i) .$$

Note that $\mathbf{F}(\mathbf{R})$ is a vector *field*.

1.2 Scalar, dot, or inner product

This is a product of two vectors results in a scalar quantity and is defined as follows for 3-component vectors:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 .$$

Note that

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2 = |\mathbf{a}|^2 = a^2.$$

The following laws of multiplication follow immediately from the definition:

- (a) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ (commutativity)
- (b) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$ (distributivity with respect to vector addition)
- (c) $(\lambda \mathbf{a}) \cdot \mathbf{b} = \lambda(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\lambda \mathbf{b})$ scalar multiple of a scalar product of two vectors

1.2.1 Geometrical interpretation of scalar product

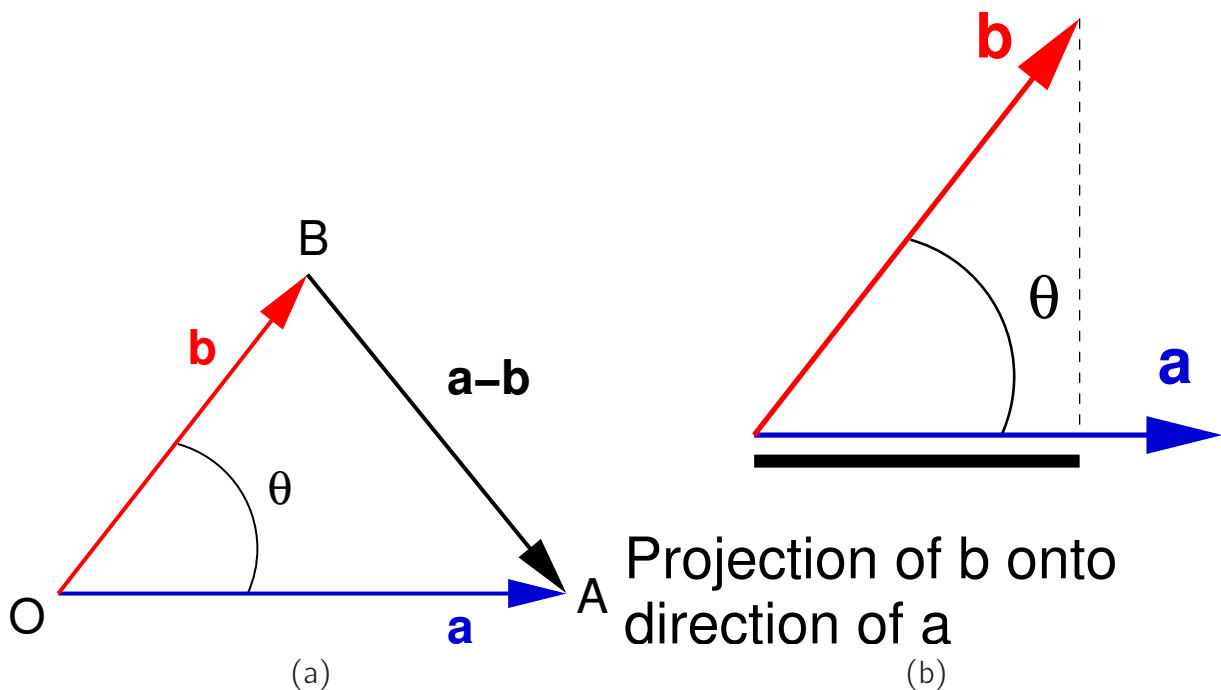


Figure 1.4: (a) Cosine rule. (b) Projection of \mathbf{b} onto \mathbf{a} .

Consider the square magnitude of the vector $\mathbf{a} - \mathbf{b}$. By the rules of the scalar product, this is

$$\begin{aligned} |\mathbf{a} - \mathbf{b}|^2 &= (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} - 2(\mathbf{a} \cdot \mathbf{b}) \\ &= a^2 + b^2 - 2(\mathbf{a} \cdot \mathbf{b}) \end{aligned}$$

But, by the cosine rule for the triangle OAB (Figure 1.4a), the length AB^2 is given by

$$|\mathbf{a} - \mathbf{b}|^2 = a^2 + b^2 - 2ab \cos \theta$$

where θ is the angle between the two vectors. It follows that

$$\mathbf{a} \cdot \mathbf{b} = ab \cos \theta,$$

which is independent of the co-ordinate system used, and that $|\mathbf{a} \cdot \mathbf{b}| \leq ab$. Conversely, the cosine of the angle between vectors \mathbf{a} and \mathbf{b} is given by $\cos \theta = \mathbf{a} \cdot \mathbf{b} / ab$.

1.2.2 Projection of one vector onto the other

Another way of describing the scalar product is as the product of the magnitude of one vector and the component of the other in the direction of the first, since $b \cos \theta$ is the component of \mathbf{b} in the direction of \mathbf{a} and vice versa (Figure 1.4b).

Projection is particularly useful when the second vector is a unit vector — $\mathbf{a} \cdot \hat{\mathbf{i}}$ is the component of \mathbf{a} in the direction of $\hat{\mathbf{i}}$.

Notice that if we wanted the **vector component** of \mathbf{b} in the direction of \mathbf{a} we would write

$$(\mathbf{b} \cdot \hat{\mathbf{a}})\hat{\mathbf{a}} = \frac{(\mathbf{b} \cdot \mathbf{a})\mathbf{a}}{a^2}.$$

In the particular case $\mathbf{a} \cdot \mathbf{b} = 0$, the angle between the two vectors is a right angle and the vectors are said to be mutually orthogonal or perpendicular — neither vector has any component in the direction of the other.

An orthonormal coordinate system is characterised by $\hat{\mathbf{i}} \cdot \hat{\mathbf{i}} = \hat{\mathbf{j}} \cdot \hat{\mathbf{j}} = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}} = 1$; and $\hat{\mathbf{i}} \cdot \hat{\mathbf{j}} = \hat{\mathbf{j}} \cdot \hat{\mathbf{k}} = \hat{\mathbf{k}} \cdot \hat{\mathbf{i}} = 0$.

1.2.3 A scalar product is an “inner product”

So far we have been writing our vectors as row vectors $\mathbf{a} = [a_1, a_2, a_3]$. This is convenient because it takes up less room than writing column vectors

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

In matrix algebra vectors are more usually defined as column vectors, as in

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

1.2. SCALAR, DOT, OR INNER PRODUCT

and a row vector is written as \mathbf{a}^\top . Now for most of our work we can be quite relaxed about this minor difference, but here let us be fussy.

Why? Simply to point out at that the scalar product is also the **inner product** more commonly used in linear algebra. Defined as $\mathbf{a}^\top \mathbf{b}$ when vectors are column vectors as

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^\top \mathbf{b} = [a_1, a_2, a_3] \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3 .$$

Here we treat a n -dimensional column vector as an $n \times 1$ matrix.

(Remember that if you multiply two matrices $\mathbf{M}_{m \times n} \mathbf{N}_{n \times p}$ then \mathbf{M} must have the same columns as \mathbf{N} has rows (here denoted by n) and the result has size (rows \times columns) of $m \times p$. So for n -dimensional column vectors \mathbf{a} and \mathbf{b} , \mathbf{a}^\top is a $1 \times n$ matrix and \mathbf{b} is $n \times 1$ matrix, so the product $\mathbf{a}^\top \mathbf{b}$ is a 1×1 matrix, which is (at last!) a scalar.)

♣ Examples

Q1. A force \mathbf{F} is applied to an object as it moves by a small amount $\delta \mathbf{r}$. What work is done on the object by the force?

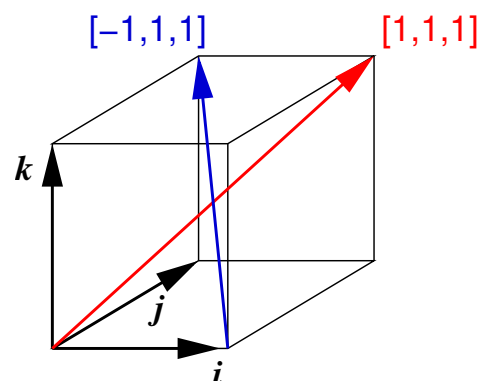
A1. The work done is equal to the component of force in the direction of the displacement multiplied by the displacement itself. This is just a scalar product:

$$\delta W = \mathbf{F} \cdot \delta \mathbf{r} .$$

Q2. A cube has four diagonals, connecting opposite vertices. What is the angle between an adjacent pair?

A2. Well, you could plod through using Pythagoras' theorem to find the length of the diagonal from cube vertex to cube centre, and perhaps you should to check the following answer.

The directions of the diagonals are $[\pm 1, \pm 1, \pm 1]$. The ones shown in the figure are $[1, 1, 1]$ and $[-1, 1, 1]$. The angle is thus



$$\theta = \cos^{-1} \frac{[1, 1, 1] \cdot [-1, 1, 1]}{\sqrt{1^2 + 1^2 + 1^2} \sqrt{-1^2 + 1^2 + 1^2}} = \cos^{-1} \frac{1}{3}$$

Q3. A pinball moving in a plane with velocity \mathbf{s} bounces (in a purely elastic impact) from a baffle whose endpoints are \mathbf{p} and \mathbf{q} . What is the velocity vector after the bounce?

A3. Best to refer everything to a coordinate frame with principal directions $\hat{\mathbf{u}}$ along and $\hat{\mathbf{v}}$ perpendicular to the baffle:

$$\hat{\mathbf{u}} = \frac{\mathbf{q} - \mathbf{p}}{|\mathbf{q} - \mathbf{p}|}$$

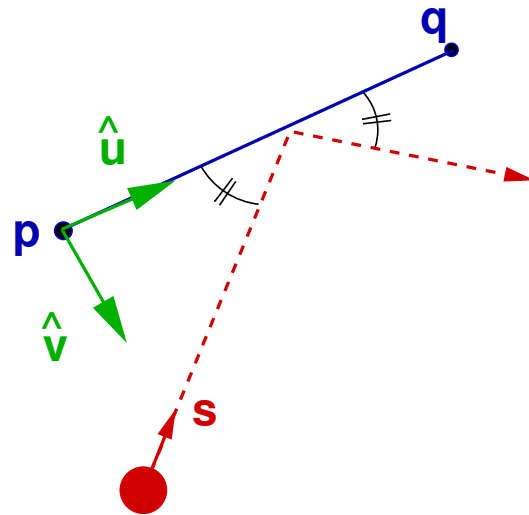
$$\hat{\mathbf{v}} = \mathbf{u}^\perp = [-u_y, u_x]$$

Thus the velocity before impact is

$$\mathbf{s}_{\text{before}} = (\mathbf{s} \cdot \hat{\mathbf{u}})\hat{\mathbf{u}} + (\mathbf{s} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}$$

After the impact, the component of velocity in the direction of the baffle is unchanged and the component normal to the baffle is negated:

$$\mathbf{s}_{\text{after}} = (\mathbf{s} \cdot \hat{\mathbf{u}})\hat{\mathbf{u}} - (\mathbf{s} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}$$



1.2.4 Direction cosines use projection

Direction cosines are commonly used in the field of crystallography. The quantities

$$\lambda = \frac{\mathbf{a} \cdot \hat{\mathbf{i}}}{a}, \quad \mu = \frac{\mathbf{a} \cdot \hat{\mathbf{j}}}{a}, \quad \nu = \frac{\mathbf{a} \cdot \hat{\mathbf{k}}}{a}$$

represent the cosines of the angles which the vector \mathbf{a} makes with the co-ordinate vectors $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ and are known as the direction cosines of the vector \mathbf{a} . Since $\mathbf{a} \cdot \hat{\mathbf{i}} = a_1$ etc, it follows immediately that $\mathbf{a} = a(\lambda\hat{\mathbf{i}} + \mu\hat{\mathbf{j}} + \nu\hat{\mathbf{k}})$ and $\lambda^2 + \mu^2 + \nu^2 = \frac{1}{a^2}[a_1^2 + a_2^2 + a_3^2] = 1$

1.3 Vector or cross product

The vector product of two vectors \mathbf{a} and \mathbf{b} is denoted by $\mathbf{a} \times \mathbf{b}$ and is defined as follows

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2)\hat{\mathbf{i}} + (a_3b_1 - a_1b_3)\hat{\mathbf{j}} + (a_1b_2 - a_2b_1)\hat{\mathbf{k}}.$$

1.3. VECTOR OR CROSS PRODUCT

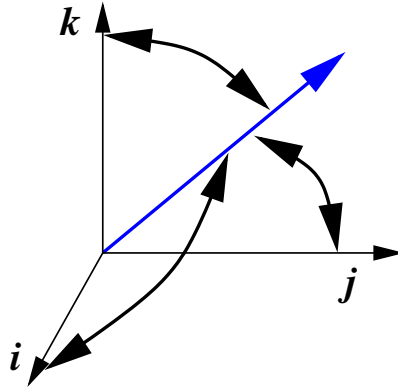


Figure 1.5: The direction cosines are cosines of the angles shown.

It is MUCH more easily remembered in terms of the (pseudo-)determinant

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

where the top row consists of the vectors \hat{i} , \hat{j} , \hat{k} rather than scalars.

Since a determinant with two equal rows has value zero, it follows that $\mathbf{a} \times \mathbf{a} = \mathbf{0}$. It is also easily verified that $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0$, so that $\mathbf{a} \times \mathbf{b}$ is orthogonal (perpendicular) to both \mathbf{a} and \mathbf{b} , as shown in Figure 1.6.

Note that $\hat{i} \times \hat{j} = \hat{k}$, $\hat{j} \times \hat{k} = \hat{i}$, and $\hat{k} \times \hat{i} = \hat{j}$.

The magnitude of the vector product can be obtained by showing that

$$|\mathbf{a} \times \mathbf{b}|^2 + (\mathbf{a} \cdot \mathbf{b})^2 = a^2 b^2$$

from which it follows that

$$|\mathbf{a} \times \mathbf{b}| = ab \sin \theta ,$$

which is again independent of the co-ordinate system used. This is left as an exercise.

Unlike the scalar product, the vector product does not satisfy commutativity but is in fact anti-commutative, in that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. Moreover the vector product does not satisfy the associative law of multiplication either since, as we shall see later $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

Since the vector product is known to be orthogonal to both the vectors which form the product, it merely remains to specify its sense with respect to these vectors. Assuming that the co-ordinate vectors form a right-handed set in the order $\hat{i}, \hat{j}, \hat{k}$ it can be seen that the sense of the the vector product is also right handed, i.e

the vector product has the same sense as the co-ordinate system used.

$$\hat{i} \times \hat{j} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \hat{k}.$$

In practice, figure out the direction from a right-handed screw twisted from the first to second vector as shown in Figure 1.6(a).

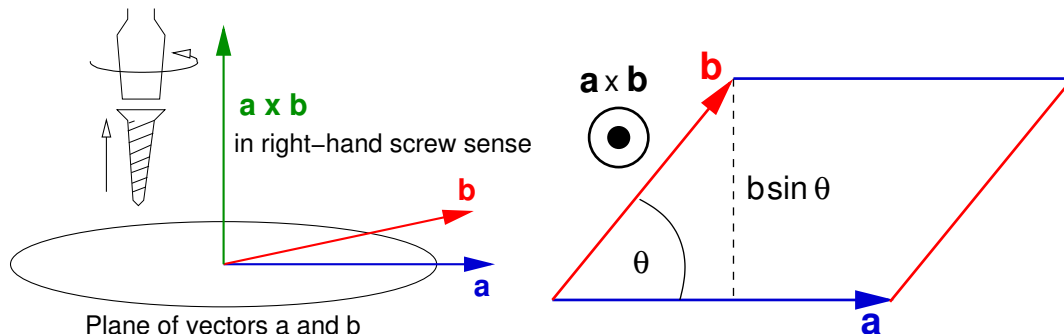


Figure 1.6: (a) The vector product is orthogonal to both **a** and **b**. Twist from first to second and move in the direction of a right-handed screw. (b) Area of parallelogram is $ab \sin \theta$.

1.3.1 Geometrical interpretation of vector product

The magnitude of the vector product ($\mathbf{a} \times \mathbf{b}$) is equal to the area of the parallelogram whose sides are parallel to, and have lengths equal to the magnitudes of, the vectors **a** and **b** (Figure 1.6b). Its direction is perpendicular to the parallelogram.

♣ Example

Q. **g** is vector from A [1,2,3] to B [3,4,5].

$\hat{\ell}$ is the unit vector in the direction from O to A.

Find $\hat{\mathbf{m}}$, a UNIT vector along $\mathbf{g} \times \hat{\ell}$

Verify that $\hat{\mathbf{m}}$ is perpendicular to $\hat{\ell}$.

Find $\hat{\mathbf{n}}$, the third member of a right-handed coordinate set $\hat{\ell}, \hat{\mathbf{m}}, \hat{\mathbf{n}}$.

A.

$$\mathbf{g} = [3, 4, 5] - [1, 2, 3] = [2, 2, 2]$$

$$\hat{\ell} = \frac{1}{\sqrt{14}}[1, 2, 3]$$

$$\mathbf{g} \times \hat{\ell} = \frac{1}{\sqrt{14}} \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 2 & 2 & 2 \\ 1 & 2 & 3 \end{vmatrix} = \frac{1}{\sqrt{14}}[2, -4, 2]$$

1.3. VECTOR OR CROSS PRODUCT

Hence

$$\hat{\mathbf{m}} = \frac{1}{\sqrt{24}}[2, -4, 2]$$

and

$$\hat{\mathbf{n}} = \hat{\boldsymbol{\ell}} \times \hat{\mathbf{m}}$$

Lecture 2

Multiple Products. Geometry using Vectors

2.1 Triple and multiple products

Using mixtures of the pairwise scalar product and vector product, it is possible to derive “triple products” between three vectors, and indeed n -products between n vectors.

There is nothing about these that you cannot work out from the definitions of pairwise scalar and vector products already given, but some have interesting geometric interpretations, so it is worth looking at these.

2.1.1 Scalar triple product

This is the scalar product of a vector product and a third vector, i.e. $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$. Using the pseudo-determinant expression for the vector product, we see that the scalar triple product can be represented as the true determinant

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

You will recall that if you swap a pair of rows of a determinant, its sign changes; hence if you swap two pairs, its sign stays the same.

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \quad \begin{matrix} \text{1st SWAP} \\ + \end{matrix} \quad \begin{vmatrix} c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{vmatrix} \quad \begin{matrix} \text{2nd SWAP} \\ - \end{matrix} \quad \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \quad \begin{matrix} + \end{matrix}$$

This says that

$$(i) \quad \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \quad (\text{Called cyclic permutation.})$$

- (ii) $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = -\mathbf{b} \cdot (\mathbf{a} \times \mathbf{c})$ and so on. (Called anti-cyclic permutation.)
- (iii) The fact that $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ allows the scalar triple product to be written as $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$. This notation is not very helpful, and we will try to avoid it below.

2.1.2 Geometrical interpretation of scalar triple product

The scalar triple product gives the volume of the parallelopiped whose sides are represented by the vectors \mathbf{a} , \mathbf{b} , and \mathbf{c} .

We saw earlier that the vector product $(\mathbf{a} \times \mathbf{b})$ has magnitude equal to the area of the base, and direction perpendicular to the base. The *component* of \mathbf{c} in this direction is equal to the height of the parallelopiped shown in Figure 2.1(a).

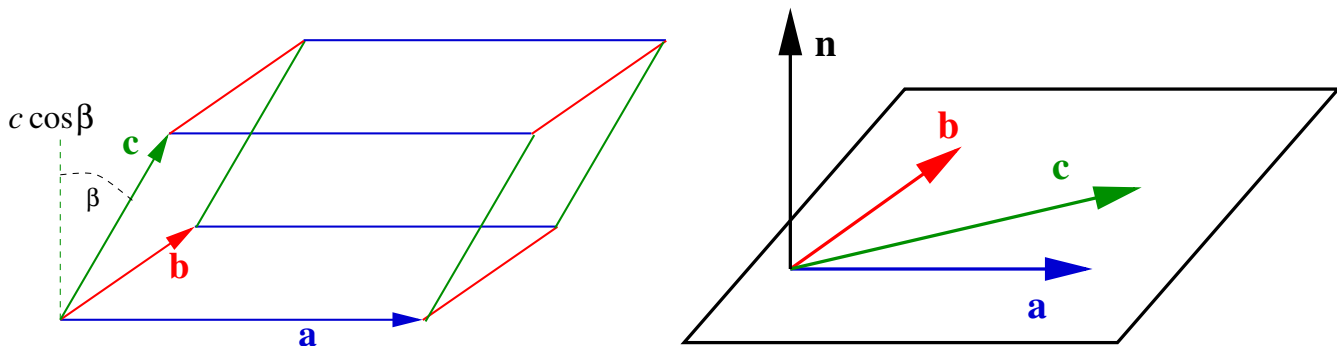


Figure 2.1: (a) Scalar triple product equals volume of parallelopiped. (b) Coplanarity yields zero scalar triple product.

2.1.3 Linearly dependent vectors

If the scalar triple product of three vectors is zero

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 0$$

then the vectors are **linearly dependent**. That is, one can be expressed as a linear combination of the others. For example,

$$\mathbf{a} = \lambda \mathbf{b} + \mu \mathbf{c}$$

where λ and μ are scalar coefficients.

You can see this immediately in two ways:

- The determinant would have one row that was a linear combination of the others. You'll remember that by doing row operations, you could reach a row of zeros, and so the determinant is zero.
- The parallelopiped would have zero volume if squashed flat. In this case all three vectors lie in a plane, and so any one is a linear combination of the other two. (Figure 2.1b.)

2.1.4 Vector triple product

This is defined as the vector product of a vector with a vector product, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. Now, the vector triple product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ must be perpendicular to $(\mathbf{b} \times \mathbf{c})$, which in turn is perpendicular to both \mathbf{b} and \mathbf{c} . Thus $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ can have no component perpendicular to \mathbf{b} and \mathbf{c} , and hence must be coplanar with them. It follows that the vector triple product must be expressible as a linear combination of \mathbf{b} and \mathbf{c} :

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \lambda \mathbf{b} + \mu \mathbf{c} .$$

The values of the coefficients can be obtained by multiplying out in component form. Only the first component need be evaluated, the others then being obtained by symmetry. That is

$$\begin{aligned} (\mathbf{a} \times (\mathbf{b} \times \mathbf{c}))_1 &= a_2(\mathbf{b} \times \mathbf{c})_3 - a_3(\mathbf{b} \times \mathbf{c})_2 \\ &= a_2(b_1c_2 - b_2c_1) + a_3(b_1c_3 - b_3c_1) \\ &= (a_2c_2 + a_3c_3)b_1 - (a_2b_2 + a_3b_3)c_1 \\ &= (a_1c_1 + a_2c_2 + a_3c_3)b_1 - (a_1b_1 + a_2b_2 + a_3b_3)c_1 \\ &= (\mathbf{a} \cdot \mathbf{c})b_1 - (\mathbf{a} \cdot \mathbf{b})c_1 \end{aligned}$$

The equivalents must be true for the 2nd and 3rd components, so we arrive at the identity

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} .$$

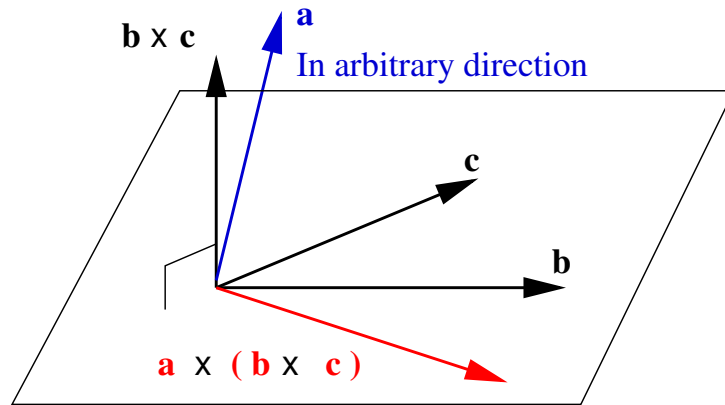


Figure 2.2: Vector triple product.

2.1.5 Projection using vector triple product

An example of the application of this formula is as follows. Suppose \mathbf{v} is a vector and we want its projection into the xy -plane. The component of \mathbf{v} in the z direction is $\mathbf{v} \cdot \hat{\mathbf{k}}$, so the projection we seek is $\mathbf{v} - (\mathbf{v} \cdot \hat{\mathbf{k}})\hat{\mathbf{k}}$. Writing $\hat{\mathbf{k}} \leftarrow \mathbf{a}$, $\mathbf{v} \leftarrow \mathbf{b}$,

$$\hat{\mathbf{k}} \leftarrow \mathbf{c},$$

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \\ \downarrow \quad \downarrow \quad \downarrow \\ \hat{\mathbf{k}} \times (\mathbf{v} \times \hat{\mathbf{k}}) &= (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}})\mathbf{v} - (\hat{\mathbf{k}} \cdot \mathbf{v})\hat{\mathbf{k}} \\ &= \mathbf{v} - (\mathbf{v} \cdot \hat{\mathbf{k}})\hat{\mathbf{k}} \end{aligned}$$

So $\mathbf{v} - (\mathbf{v} \cdot \hat{\mathbf{k}})\hat{\mathbf{k}} = \hat{\mathbf{k}} \times (\mathbf{v} \times \hat{\mathbf{k}})$.

(Hot stuff! But the expression $\mathbf{v} - (\mathbf{v} \cdot \hat{\mathbf{k}})\hat{\mathbf{k}}$ is much easier to understand, and cheaper to compute!)

2.1.6 Other repeated products

Many combinations of vector and scalar products are possible, but we consider only one more, namely the vector quadruple product $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})$. By regarding $\mathbf{a} \times \mathbf{b}$ as a single vector, we see that this vector must be representable as a linear combination of \mathbf{c} and \mathbf{d} . On the other hand, regarding $\mathbf{c} \times \mathbf{d}$ as a single vector, we see that it must also be a linear combination of \mathbf{a} and \mathbf{b} . This provides a means of expressing one of the vectors, say \mathbf{d} , as linear combination of the other three, as follows:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}]\mathbf{c} - [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]\mathbf{d} \\ &= [(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{a}]\mathbf{b} - [(\mathbf{c} \times \mathbf{d}) \cdot \mathbf{b}]\mathbf{a} \end{aligned}$$

Hence

$$[(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]\mathbf{d} = [(\mathbf{b} \times \mathbf{c}) \cdot \mathbf{d}]\mathbf{a} + [(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{d}]\mathbf{b} + [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}]\mathbf{c}$$

or

$$\mathbf{d} = \frac{[(\mathbf{b} \times \mathbf{c}) \cdot \mathbf{d}]\mathbf{a} + [(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{d}]\mathbf{b} + [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}]\mathbf{c}}{[(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}.$$

This is not something to remember off by heart, but it is worth remembering that the projection of a vector on any arbitrary basis set is unique.

♣ Example

Q1 Use the quadruple vector product to express the vector $\mathbf{d} = [3, 2, 1]$ in terms of the vectors $\mathbf{a} = [1, 2, 3]$, $\mathbf{b} = [2, 3, 1]$ and $\mathbf{c} = [3, 1, 2]$.

A1 Grinding away at the determinants, we find

$$[(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}] = -18; \quad [(\mathbf{b} \times \mathbf{c}) \cdot \mathbf{d}] = 6; \quad [(\mathbf{c} \times \mathbf{a}) \cdot \mathbf{d}] = -12; \quad [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}] = -12$$

$$\text{So, } \mathbf{d} = (-\mathbf{a} + 2\mathbf{b} + 2\mathbf{c})/3.$$

2.2. GEOMETRY USING VECTORS: LINES, PLANES

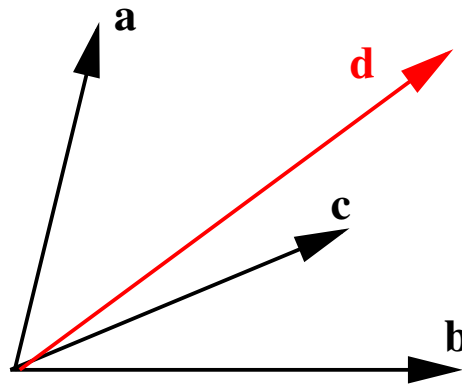


Figure 2.3: The projection of a (3-)vector onto a set of (3) basis vectors is unique. I.e. in $\mathbf{d} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}$, the set $\{\alpha, \beta, \gamma\}$ is unique.

2.2 Geometry using vectors: lines, planes

2.2.1 The equation of a line

The equation of the line passing through the point whose position vector is \mathbf{a} and lying in the direction of vector \mathbf{b} is

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$$

where λ is a scalar parameter. If you make \mathbf{b} a unit vector, $\mathbf{r} = \mathbf{a} + \lambda\hat{\mathbf{b}}$ then λ will represent metric length.

For a line defined by two points \mathbf{a}_1 and \mathbf{a}_2

$$\mathbf{r} = \mathbf{a}_1 + \lambda(\mathbf{a}_2 - \mathbf{a}_1)$$

or for the unit version

$$\mathbf{r} = \mathbf{a}_1 + \lambda \frac{(\mathbf{a}_2 - \mathbf{a}_1)}{|\mathbf{a}_2 - \mathbf{a}_1|}$$

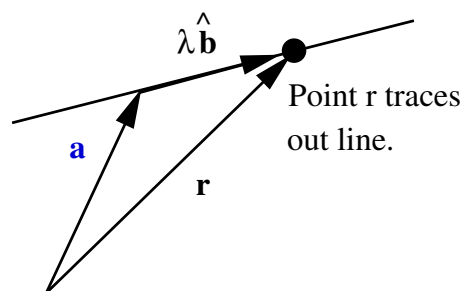


Figure 2.4: Equation of a line. With $\hat{\mathbf{b}}$ a unit vector, λ is in the length units established by the definition of \mathbf{a} .

2.2.2 The shortest distance from a point to a line

Referring to Figure 2.5(a) the vector \mathbf{p} from \mathbf{c} to any point on the line is $\mathbf{p} = \mathbf{a} + \lambda \hat{\mathbf{b}} - \mathbf{c} = (\mathbf{a} - \mathbf{c}) + \lambda \hat{\mathbf{b}}$ which has length squared $p^2 = (\mathbf{a} - \mathbf{c})^2 + \lambda^2 + 2\lambda(\mathbf{a} - \mathbf{c}) \cdot \hat{\mathbf{b}}$. Rather than minimizing length, it is easier to minimize length-squared. The minimum is found when $d p^2 / d\lambda = 0$, ie when

$$\lambda = -(\mathbf{a} - \mathbf{c}) \cdot \hat{\mathbf{b}}.$$

So the minimum length vector is

$$\mathbf{p} = (\mathbf{a} - \mathbf{c}) - ((\mathbf{a} - \mathbf{c}) \cdot \hat{\mathbf{b}})\hat{\mathbf{b}}.$$

You might spot that is the component of $(\mathbf{a} - \mathbf{c})$ perpendicular to $\hat{\mathbf{b}}$ (as expected!). Furthermore, using the result of Section 2.1.5,

$$\mathbf{p} = \hat{\mathbf{b}} \times [(\mathbf{a} - \mathbf{c}) \times \hat{\mathbf{b}}].$$

Because $\hat{\mathbf{b}}$ is a unit vector, and is orthogonal to $[(\mathbf{a} - \mathbf{c}) \times \hat{\mathbf{b}}]$, the modulus of the vector can be written rather more simply as just

$$\rho_{\min} = |(\mathbf{a} - \mathbf{c}) \times \hat{\mathbf{b}}|.$$

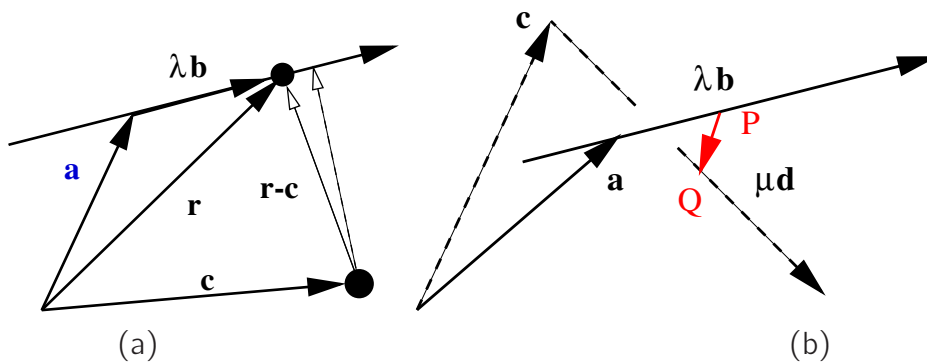


Figure 2.5: (a) Shortest distance point to line. (b) Shortest distance, line to line.

2.2.3 The shortest distance between two straight lines

If the shortest distance between a point and a line is along the perpendicular, then the shortest distance between the two straight lines $\mathbf{r} = \mathbf{a} + \lambda \hat{\mathbf{b}}$ and $\mathbf{r} = \mathbf{c} + \mu \hat{\mathbf{d}}$ must be found as the length of the vector which is mutually perpendicular to the lines.

The unit vector along the mutual perpendicular is

$$\hat{\mathbf{p}} = (\hat{\mathbf{b}} \times \hat{\mathbf{d}}) / |\hat{\mathbf{b}} \times \hat{\mathbf{d}}|.$$

(Yes, don't forget that $\hat{\mathbf{b}} \times \hat{\mathbf{d}}$ is NOT a unit vector. $\hat{\mathbf{b}}$ and $\hat{\mathbf{d}}$ are not orthogonal, so there is a $\sin \theta$ lurking!)

The minimum length is therefore the component of $\mathbf{a} - \mathbf{c}$ in this direction

$$\rho_{\min} = |(\mathbf{a} - \mathbf{c}) \cdot (\hat{\mathbf{b}} \times \hat{\mathbf{d}}) / |\hat{\mathbf{b}} \times \hat{\mathbf{d}}||.$$

♣ Example

- Q** Two long straight pipes are specified using Cartesian co-ordinates as follows:
 Pipe A has diameter 0.8 and its axis passes through points (2, 5, 3) and (7, 10, 8).
 Pipe B has diameter 1.0 and its axis passes through the points (0, 6, 3) and (−12, 0, 9).
 Determine whether the pipes need to be realigned to avoid intersection.

- A** Each pipe axis is defined using two points. The vector equation of the axis of pipe A is

$$\mathbf{r} = [2, 5, 3] + \lambda'[5, 5, 5] = [2, 5, 3] + \lambda[1, 1, 1]/\sqrt{3}$$

The equation of the axis of pipe B is

$$\mathbf{r} = [0, 6, 3] + \mu'[12, 6, 6] = [0, 6, 3] + \mu[-2, -1, 1]/\sqrt{6}$$

The perpendicular to the two axes has direction

$$[1, 1, 1] \times [-2, -1, 1] = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 1 & 1 \\ -2 & -1 & 1 \end{vmatrix} = [2, -3, 1] = \mathbf{p}$$

The length of the mutual perpendicular is

$$(\mathbf{a} - \mathbf{c}) \cdot \frac{[2, -3, 1]}{\sqrt{14}} = [2, -1, 0] \cdot \frac{[2, -3, 1]}{\sqrt{14}} = 1.87.$$

But the sum of the radii of the two pipes is $0.4 + 0.5 = 0.9$. Hence the pipes do not intersect.

2.2.4 The equation of a plane

There are a number of ways of specifying the equation of a plane.

1. If \mathbf{b} and \mathbf{c} are two non-parallel vectors (ie $\mathbf{b} \times \mathbf{c} \neq \mathbf{0}$), then the equation of the plane passing through the point \mathbf{a} and parallel to the vectors \mathbf{b} and \mathbf{c} may be written in the form

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{b} + \mu\mathbf{c}$$

where λ, μ are scalar parameters. Note that \mathbf{b} and \mathbf{c} are free vectors, so don't have to lie in the plane (Figure 2.6(a).)

2. Figure 2.6(b) shows the plane defined by three non-collinear points \mathbf{a} , \mathbf{b} and \mathbf{c} in the plane (note that the vectors \mathbf{b} and \mathbf{c} are position vectors, not free vectors as in the previous case). The equation might be written as

$$\mathbf{r} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) + \mu(\mathbf{c} - \mathbf{a})$$

3. Figure 2.6(c) illustrates another description is in terms of the unit normal to the plane $\hat{\mathbf{n}}$ and a point \mathbf{a} in the plane

$$\mathbf{r} \cdot \hat{\mathbf{n}} = \mathbf{a} \cdot \hat{\mathbf{n}}.$$

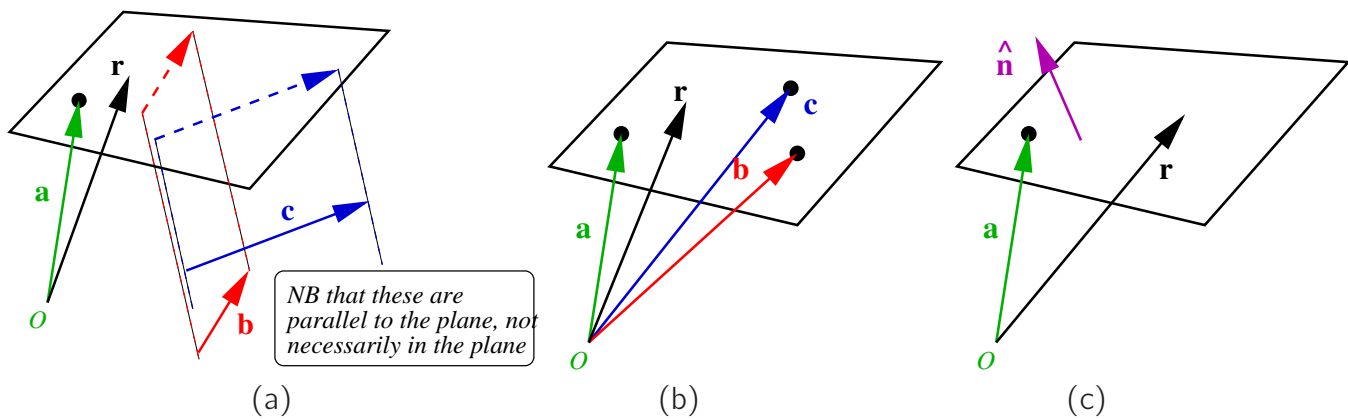


Figure 2.6: (a) Plane defined using point and two lines. (b) Plane defined using three points. (c) Plane defined using point and normal. Vector \mathbf{r} is the position vector of a general point in the plane.

2.2.5 The shortest distance from a point to a plane

The shortest distance from a point \mathbf{d} to the plane is along the perpendicular. Depending on how the plane is defined, this can be written as

$$D = |(\mathbf{d} - \mathbf{a}) \cdot \hat{\mathbf{n}}| \quad \text{or} \quad D = \frac{|(\mathbf{d} - \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{c})|}{|\mathbf{b} \times \mathbf{c}|}.$$

2.3 Solution of vector equations

It is sometimes required to obtain the most general vector which satisfies a given vector relationship. This is very much like obtaining the locus of a point. The best method of proceeding in such a case is as follows:

- (i) Decide upon a system of three co-ordinate vectors using two non-parallel vectors appearing in the vector relationship. These might be \mathbf{a} , \mathbf{b} and their vector product $(\mathbf{a} \times \mathbf{b})$.
- (ii) Express the unknown vector \mathbf{x} as a linear combination of these vectors

$$\mathbf{x} = \lambda \mathbf{a} + \mu \mathbf{b} + \nu \mathbf{a} \times \mathbf{b}$$

where λ, μ, ν are scalars to be found.

- (iii) Substitute the above expression for \mathbf{x} into the vector relationship to determine the constraints on λ, μ and ν for the relationship to be satisfied.

♣ Example

Q Solve the vector equation $\mathbf{x} = \mathbf{x} \times \mathbf{a} + \mathbf{b}$.

A Step (i): Set up basis vectors \mathbf{a} , \mathbf{b} and their vector product $\mathbf{a} \times \mathbf{b}$.

Step (ii): $\mathbf{x} = \lambda \mathbf{a} + \mu \mathbf{b} + \nu \mathbf{a} \times \mathbf{b}$.

2.4. ROTATION, ANGULAR VELOCITY/ACCELERATION AND MOMENTS

Step (iii): Bung this expression for \mathbf{x} into the equation!

$$\begin{aligned}\lambda \mathbf{a} + \mu \mathbf{b} + \nu \mathbf{a} \times \mathbf{b} &= (\lambda \mathbf{a} + \mu \mathbf{b} + \nu \mathbf{a} \times \mathbf{b}) \times \mathbf{a} + \mathbf{b} \\ &= \mathbf{0} + \mu(\mathbf{b} \times \mathbf{a}) + \nu(\mathbf{a} \times \mathbf{b}) \times \mathbf{a} + \mathbf{b} \\ &= -\nu(\mathbf{a} \cdot \mathbf{b})\mathbf{a} + (\nu a^2 + 1)\mathbf{b} - \mu(\mathbf{a} \times \mathbf{b})\end{aligned}$$

We have learned that any vector has a unique expression in terms of a basis set, so that the coefficients of \mathbf{a} , \mathbf{b} and $\mathbf{a} \times \mathbf{b}$ on either side of the equation must be equal.

$$\begin{aligned}\Rightarrow \lambda &= -\nu(\mathbf{a} \cdot \mathbf{b}) \\ \mu &= \nu a^2 + 1 \\ \nu &= -\mu\end{aligned}$$

so that

$$\mu = \frac{1}{1 + a^2} \quad \nu = -\frac{1}{1 + a^2} \quad \lambda = \frac{\mathbf{a} \cdot \mathbf{b}}{1 + a^2}.$$

So finally the solution is the single point:

$$\mathbf{x} = \frac{1}{1 + a^2} ((\mathbf{a} \cdot \mathbf{b})\mathbf{a} + \mathbf{b} - (\mathbf{a} \times \mathbf{b}))$$

2.4 Rotation, angular velocity/acceleration and moments

A rotation can be represented by a vector whose direction is along the axis of rotation in the sense of a r-h screw, and whose magnitude is proportional to the size of the rotation (Fig. 2.7). The same idea can be extended to the derivatives, that is, angular velocity $\boldsymbol{\omega}$ and angular acceleration $\dot{\boldsymbol{\omega}}$.

Angular accelerations arise because of a moment (or torque) on a body. In mechanics, the moment of a force \mathbf{F} about a point Q is defined to have magnitude $M = Fd$, where d is the perpendicular distance between Q and the line of action L of \mathbf{F} .

The vector equation for moment is

$$\mathbf{M} = \mathbf{r} \times \mathbf{F}$$

where \mathbf{r} is the vector from Q to *any* point on the line of action L of force \mathbf{F} . The resulting angular acceleration vector is in the same direction as the moment vector.

The instantaneous velocity of any point P on a rigid body undergoing pure rotation can be defined by a vector product as follows. The angular velocity vector $\boldsymbol{\omega}$ has

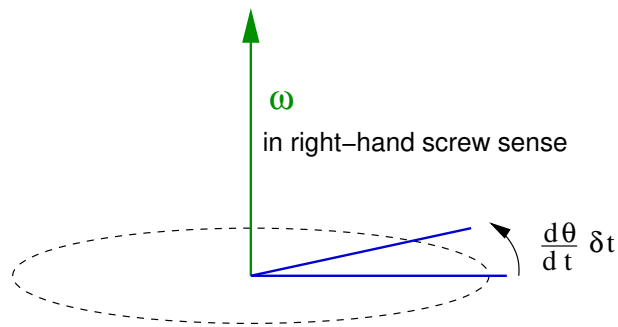


Figure 2.7: The angular velocity vector $\boldsymbol{\omega}$ is along the axis of rotation and has magnitude equal to the rate of rotation.

magnitude equal to the angular speed of rotation of the body and with direction the same as that of the r-h screw. If \mathbf{r} is the vector OP , where the origin O can be taken to be any point on the axis of rotation, then the velocity \mathbf{v} of P due to the rotation is given, in both magnitude and direction, by the vector product

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}.$$

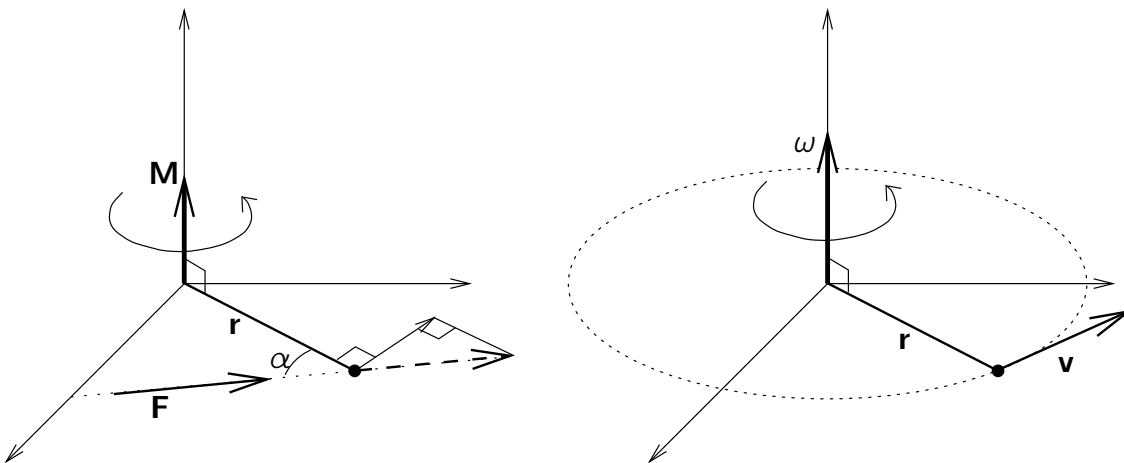


Figure 2.8:

Lecture 3

Differentiating Vector Functions of a Single Variable

Your experience of differentiation and integration has extended as far as *scalar* functions of single and multiple variables — functions like $f(x)$ and $f(x, y, t)$.

It should be no great surprise that we often wish to differentiate vector functions. For example, suppose you were driving along a wiggly road with position $\mathbf{r}(t)$ at time t . Differentiating $\mathbf{r}(t)$ wrt time should yield your velocity $\mathbf{v}(t)$, and differentiating $\mathbf{v}(t)$ should yield your acceleration. Let's see how to do this.

3.1 Differentiation of a vector

The derivative of a vector function $\mathbf{a}(p)$ of a single parameter p is

$$\mathbf{a}'(p) = \lim_{\delta p \rightarrow 0} \frac{\mathbf{a}(p + \delta p) - \mathbf{a}(p)}{\delta p} .$$

If we write \mathbf{a} in terms of components relative to a FIXED coordinate system ($\hat{i}, \hat{j}, \hat{k}$ constant)

$$\mathbf{a}(p) = a_1(p)\hat{i} + a_2(p)\hat{j} + a_3(p)\hat{k}$$

then

$$\mathbf{a}'(p) = \frac{da_1}{dp}\hat{i} + \frac{da_2}{dp}\hat{j} + \frac{da_3}{dp}\hat{k} .$$

That is, in order to differentiate a vector function, one simply differentiates each component separately. This means that all the familiar rules of differentiation apply, and they don't get altered by vector operations like scalar product and vector products.

Thus, for example:

$$\frac{d}{dp}(\mathbf{a} \times \mathbf{b}) = \frac{d\mathbf{a}}{dp} \times \mathbf{b} + \mathbf{a} \times \frac{d\mathbf{b}}{dp}$$

$$\frac{d}{dp}(\mathbf{a} \cdot \mathbf{b}) = \frac{d\mathbf{a}}{dp} \cdot \mathbf{b} + \mathbf{a} \cdot \frac{d\mathbf{b}}{dp} .$$

Note that $d\mathbf{a}/dp$ has a different direction and a different magnitude from \mathbf{a} .

Likewise, as you might expect, the chain rule still applies. If $\mathbf{a} = \mathbf{a}(u)$ and $u = u(t)$, say:

$$\frac{d}{dt}\mathbf{a} = \frac{d\mathbf{a}}{du} \frac{du}{dt}$$

♣ Examples

Q A 3D vector \mathbf{a} of constant magnitude is varying over time. What can you say about the direction of $\dot{\mathbf{a}}$?

A Using intuition: if only the direction is changing, then the vector must be tracing out points on the surface of a sphere. We would guess that the derivative $\dot{\mathbf{a}}$ is orthogonal to \mathbf{a} .

To prove this write

$$\frac{d}{dt}(\mathbf{a} \cdot \mathbf{a}) = \mathbf{a} \cdot \frac{d\mathbf{a}}{dt} + \frac{d\mathbf{a}}{dt} \cdot \mathbf{a} = 2\mathbf{a} \cdot \frac{d\mathbf{a}}{dt}.$$

But $(\mathbf{a} \cdot \mathbf{a}) = a^2$ which we are told is constant. So

$$\frac{d}{dt}(\mathbf{a} \cdot \mathbf{a}) = 0 \quad \Rightarrow \quad 2\mathbf{a} \cdot \frac{d\mathbf{a}}{dt} = 0$$

and hence \mathbf{a} and $d\mathbf{a}/dt$ must be perpendicular.

Q The position of a vehicle is $\mathbf{r}(u)$ where u is the amount of fuel consumed by some time t . Write down an expression for the acceleration.

A The velocity is

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{du} \frac{du}{dt}$$

$$\mathbf{a} = \frac{d}{dt} \frac{d\mathbf{r}}{dt} = \frac{d^2\mathbf{r}}{du^2} \left(\frac{du}{dt} \right)^2 + \frac{d\mathbf{r}}{du} \frac{d^2u}{dt^2}$$

3.1.1 Geometrical interpretation of vector derivatives

Let $\mathbf{r}(p)$ be a position vector tracing a space curve as some parameter p varies. The vector $\delta\mathbf{r}$ is a secant to the curve, and $\delta\mathbf{r}/\delta p$ lies in the same direction. (See Fig. 3.1.) In the limit as δp tends to zero $\delta\mathbf{r}/\delta p = d\mathbf{r}/dp$ becomes a tangent to the space curve. If the magnitude of this vector is 1 (i.e. a unit tangent), then

3.1. DIFFERENTIATION OF A VECTOR

$|d\mathbf{r}| = dp$ so the parameter p is arc-length (metric distance). More generally, however, p will not be arc-length and we will have:

$$\frac{d\mathbf{r}}{dp} = \frac{d\mathbf{r}}{ds} \frac{ds}{dp}$$

So, the direction of the derivative is that of a **tangent to the curve**, and its magnitude is $|ds/dp|$, **the rate of change of arc length w.r.t the parameter**. Of course if that parameter p is time, the magnitude $|d\mathbf{r}/dt|$ is the speed.

♣ Example

Q Draw the curve

$$\mathbf{r} = a \cos\left(\frac{s}{\sqrt{a^2 + h^2}}\right) \hat{\mathbf{i}} + a \sin\left(\frac{s}{\sqrt{a^2 + h^2}}\right) \hat{\mathbf{j}} + \frac{hs}{\sqrt{a^2 + h^2}} \hat{\mathbf{k}}$$

where s is arc length and h, a are constants. Show that the tangent $d\mathbf{r}/ds$ to the curve has a constant elevation angle w.r.t the xy -plane, and determine its magnitude.

A

$$\frac{d\mathbf{r}}{ds} = -\frac{a}{\sqrt{a^2 + h^2}} \sin\left(\frac{s}{\sqrt{a^2 + h^2}}\right) \hat{\mathbf{i}} + \frac{a}{\sqrt{a^2 + h^2}} \cos\left(\frac{s}{\sqrt{a^2 + h^2}}\right) \hat{\mathbf{j}} + \frac{h}{\sqrt{a^2 + h^2}} \hat{\mathbf{k}}$$

The projection on the xy plane has magnitude $a/\sqrt{a^2 + h^2}$ and in the z direction $h/\sqrt{a^2 + h^2}$, so the elevation angle is a constant, $\tan^{-1}(h/a)$.

We are expecting $d\mathbf{r}/ds = 1$, and indeed

$$\sqrt{a^2 \sin^2\left(\frac{s}{\sqrt{a^2 + h^2}}\right) + a^2 \cos^2\left(\frac{s}{\sqrt{a^2 + h^2}}\right) + h^2/\sqrt{a^2 + h^2}} = 1.$$

3.1.2 Arc length is a special parameter!

It might seem that we can be completely relaxed about saying that any old parameter p is arc length, but this is not the case. Why not? The reason is that arc length is special is that, whatever the parameter p ,

$$s = \int_{p_0}^p \left| \frac{d\mathbf{r}}{dp} \right| dp .$$

Perhaps another way to grasp the significance of this is using Pythagoras' theorem on a short piece of curve: in the limit as dx etc tend to zero,

$$ds^2 = dx^2 + dy^2 + dz^2 .$$

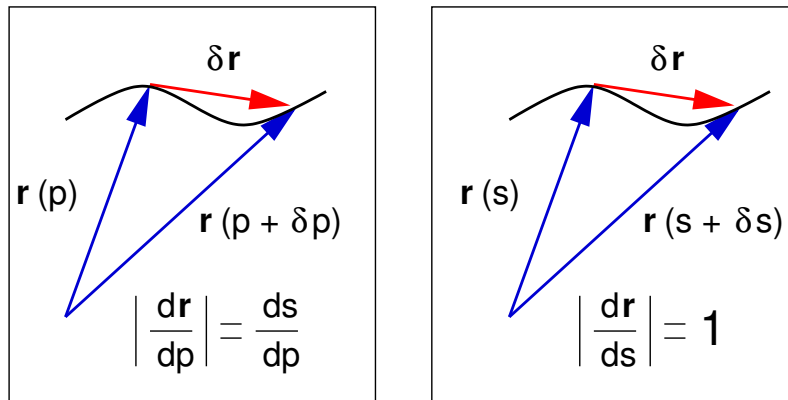


Figure 3.1: Left: $\delta \mathbf{r}$ is a secant to the curve but, in the limit as $\delta p \rightarrow 0$, becomes a tangent. Right: if the parameter is arc length s , then $|\mathbf{dr}| = ds$.

So if a curve is parameterized in terms of p

$$\frac{ds}{dp} = \sqrt{\frac{dx^2}{dp} + \frac{dy^2}{dp} + \frac{dz^2}{dp}}.$$

As an example, suppose in our earlier example we had parameterized our helix as

$$\mathbf{r} = a \cos p \hat{\mathbf{i}} + a \sin p \hat{\mathbf{j}} + hp \hat{\mathbf{k}}$$

It would be easy just to say that p was arclength, but it would not be correct because

$$\begin{aligned} \frac{ds}{dp} &= \sqrt{\frac{dx^2}{dp} + \frac{dy^2}{dp} + \frac{dz^2}{dp}} \\ &= \sqrt{a^2 \sin^2 p + a^2 \cos^2 p + h^2} = \sqrt{a^2 + h^2} \end{aligned}$$

If p really was arclength, $ds/dp = 1$. So $p/\sqrt{a^2 + h^2}$ **is** arclength, not p .

3.2 Integration of a vector function

The integration of a vector function of a single scalar variable can be regarded simply as the reverse of differentiation. In other words

$$\int_{p_1}^{p_2} \frac{d\mathbf{a}(p)}{dp} dp$$

For example the integral of the acceleration vector of a point over an interval of time is equal to the change in the velocity vector during the same time interval. However, many other, more interesting and useful, types of integral are possible, especially when the vector is a function of more than one variable. This requires the introduction of the concepts of scalar and vector fields. See later!

3.3 Curves in 3 dimensions

In the examples above, parameter p has been either arc length s or time t . It doesn't have to be, but these are the main two of interest. Later we shall look at some important results when differentiating w.r.t. time, but now let us look more closely at 3D curves defined in terms of arc length, s .

Take a piece of wire, and bend it into some arbitrary non-planar curve. This is a *space curve*. We can specify a point on the wire by specifying $\mathbf{r}(s)$ as a function of distance or arc length s along the wire.

3.3.1 The Frénet-Serret relationships

We are now going to introduce a local orthogonal coordinate frame for each point s along the curve, ie one with its origin at $\mathbf{r}(s)$. To specify a coordinate frame we need three mutually perpendicular directions, and these should be *intrinsic* to the curve, not fixed in an external reference frame. The ideas were first suggested by two French mathematicians, F-J. Frénet and J. A. Serret.

1. Tangent $\hat{\mathbf{t}}$

There is an obvious choice for the first direction at the point $\mathbf{r}(s)$, namely the **unit tangent** $\hat{\mathbf{t}}$. We already know that

$$\hat{\mathbf{t}} = \frac{d\mathbf{r}(s)}{ds}$$

2. Principal Normal $\hat{\mathbf{n}}$

Recall that earlier we proved that if \mathbf{a} was a vector of constant magnitude that varies in direction over time then $d\mathbf{a}/dt$ was perpendicular to it. Because $\hat{\mathbf{t}}$ has constant magnitude but varies over s , $d\hat{\mathbf{t}}/ds$ must be perpendicular to $\hat{\mathbf{t}}$.

Hence the principal normal $\hat{\mathbf{n}}$ is

$$\frac{d\hat{\mathbf{t}}}{ds} = \kappa \hat{\mathbf{n}} : \text{ where } \kappa \geq 0 .$$

κ is the **curvature**, and $\kappa = 0$ for a straight line. The plane containing $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$ is called the **osculating plane**.

3. The Binormal $\hat{\mathbf{b}}$

The local coordinate frame is completed by defining the binormal

$$\hat{\mathbf{b}}(s) = \hat{\mathbf{t}}(s) \times \hat{\mathbf{n}}(s) .$$

Since $\hat{\mathbf{b}} \cdot \hat{\mathbf{t}} = 0$,

$$\frac{d\hat{\mathbf{b}}}{ds} \cdot \hat{\mathbf{t}} + \hat{\mathbf{b}} \cdot \frac{d\hat{\mathbf{t}}}{ds} = \frac{d\hat{\mathbf{b}}}{ds} \cdot \hat{\mathbf{t}} + \hat{\mathbf{b}} \cdot \kappa \hat{\mathbf{n}} = 0$$

from which

$$\frac{d\hat{\mathbf{b}}}{ds} \cdot \hat{\mathbf{t}} = 0.$$

But this means that $d\hat{\mathbf{b}}/ds$ is along the direction of $\hat{\mathbf{n}}$, or

$$\frac{d\hat{\mathbf{b}}}{ds} = -\tau(s)\hat{\mathbf{n}}(s)$$

where τ is the **torsion**, and the negative sign is a matter of convention.

Differentiating $\hat{\mathbf{n}} \cdot \hat{\mathbf{t}} = 0$ and $\hat{\mathbf{n}} \cdot \hat{\mathbf{b}} = 0$, we find

$$\frac{d\hat{\mathbf{n}}}{ds} = -\kappa(s)\hat{\mathbf{t}}(s) + \tau(s)\hat{\mathbf{b}}(s).$$

The Frénet-Serret relationships:

$$d\hat{\mathbf{t}}/ds = \kappa\hat{\mathbf{n}}$$

$$d\hat{\mathbf{n}}/ds = -\kappa(s)\hat{\mathbf{t}}(s) + \tau(s)\hat{\mathbf{b}}(s)$$

$$d\hat{\mathbf{b}}/ds = -\tau(s)\hat{\mathbf{n}}(s)$$

♣ Example

Q Derive $\kappa(s)$ and $\tau(s)$ for the helix

$$\mathbf{r}(s) = a \cos\left(\frac{s}{\beta}\right) \hat{\mathbf{i}} + a \sin\left(\frac{s}{\beta}\right) \hat{\mathbf{j}} + h \left(\frac{s}{\beta}\right) \hat{\mathbf{k}}; \quad \beta = \sqrt{a^2 + h^2}$$

and comment on their values.

A We found the unit tangent earlier as

$$\hat{\mathbf{t}} = \frac{d\mathbf{r}}{ds} = \left[-\frac{a}{\beta} \sin\left(\frac{s}{\beta}\right), \frac{a}{\beta} \cos\left(\frac{s}{\beta}\right), \frac{h}{\beta} \right].$$

Differentiation gives

$$\kappa\hat{\mathbf{n}} = \frac{d\hat{\mathbf{t}}}{ds} = \left[-\frac{a}{\beta^2} \cos\left(\frac{s}{\beta}\right), -\frac{a}{\beta^2} \sin\left(\frac{s}{\beta}\right), 0 \right]$$

3.4. RADIAL AND TANGENTIAL COMPONENTS IN PLANE POLARS

Curvature is always positive, so

$$\kappa = \frac{a}{\beta^2} \quad \hat{\mathbf{n}} = \left[-\cos\left(\frac{s}{\beta}\right), -\sin\left(\frac{s}{\beta}\right), 0 \right] .$$

So the curvature is constant, and the normal is parallel to the xy -plane.

Now use

$$\hat{\mathbf{b}} = \hat{\mathbf{t}} \times \hat{\mathbf{n}} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ (-a/\beta)S & (a/\beta)C & (h/\beta) \\ -C & -S & 0 \end{vmatrix} = \left[\frac{h}{\beta} \sin\left(\frac{s}{\beta}\right), -\frac{h}{\beta} \cos\left(\frac{s}{\beta}\right), \frac{a}{\beta} \right]$$

and differentiate $\hat{\mathbf{b}}$ to find an expression for the torsion

$$\frac{d\hat{\mathbf{b}}}{ds} = \left[\frac{h}{\beta^2} \cos\left(\frac{s}{\beta}\right), \frac{h}{\beta^2} \sin\left(\frac{s}{\beta}\right), 0 \right] = \frac{-h}{\beta^2} \hat{\mathbf{n}}$$

so the torsion is

$$\tau = \frac{h}{\beta^2}$$

again a constant.

3.4 Radial and tangential components in plane polars

In plane polar coordinates, the radius vector of any point P is given by

$$\begin{aligned} \mathbf{r} &= r \cos \theta \hat{\mathbf{i}} + r \sin \theta \hat{\mathbf{j}} \\ &= r \hat{\mathbf{e}}_r \end{aligned}$$

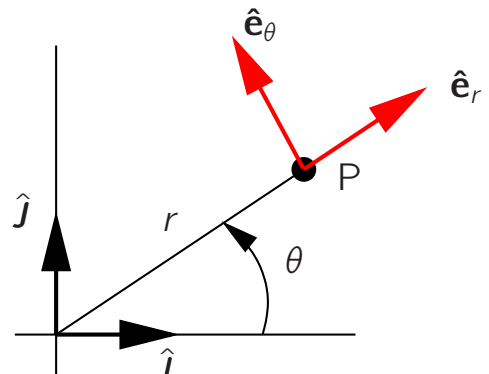
where we have introduced the unit radial vector

$$\hat{\mathbf{e}}_r = \cos \theta \hat{\mathbf{i}} + \sin \theta \hat{\mathbf{j}} .$$

The other “natural” (we’ll see why in a later lecture) unit vector in plane polars is orthogonal to $\hat{\mathbf{e}}_r$ and is

$$\hat{\mathbf{e}}_\theta = -\sin \theta \hat{\mathbf{i}} + \cos \theta \hat{\mathbf{j}}$$

so that $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_r = \hat{\mathbf{e}}_\theta \cdot \hat{\mathbf{e}}_\theta = 1$ and $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_\theta = 0$.



Now suppose P is moving so that \mathbf{r} is a function of time t . Its velocity is

$$\begin{aligned}\dot{\mathbf{r}} &= \frac{d}{dt}(r\hat{\mathbf{e}}_r) = \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\hat{\mathbf{e}}_r}{dt} \\ &= \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\theta}{dt}(-\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}}) \\ &= \frac{dr}{dt}\hat{\mathbf{e}}_r + r\frac{d\theta}{dt}\hat{\mathbf{e}}_\theta \\ &= \text{radial} + \text{tangential}\end{aligned}$$

The radial and tangential components of velocity of P are therefore dr/dt and $rd\theta/dt$, respectively.

Differentiating a second time gives the acceleration of P

$$\begin{aligned}\ddot{\mathbf{r}} &= \frac{d^2r}{dt^2}\hat{\mathbf{e}}_r + \frac{dr}{dt}\frac{d\theta}{dt}\hat{\mathbf{e}}_\theta + \frac{dr}{dt}\frac{d\theta}{dt}\hat{\mathbf{e}}_\theta + r\frac{d^2\theta}{dt^2}\hat{\mathbf{e}}_\theta - r\frac{d\theta}{dt}\frac{d\theta}{dt}\hat{\mathbf{e}}_r \\ &= \left[\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right]\hat{\mathbf{e}}_r + \left[2\frac{dr}{dt}\frac{d\theta}{dt} + r\frac{d^2\theta}{dt^2}\right]\hat{\mathbf{e}}_\theta\end{aligned}$$

3.5 Rotating systems

Consider a body which is rotating with constant angular velocity $\boldsymbol{\omega}$ about some axis passing through the origin. Assume the origin is fixed, and that we are sitting in a fixed coordinate system $Oxyz$.

If $\boldsymbol{\rho}$ is a vector of constant magnitude and constant direction in the rotating system, then its representation \mathbf{r} in the fixed system must be a function of t .

$$\mathbf{r}(t) = \mathbf{R}(t)\boldsymbol{\rho}$$

At any instant as observed in the fixed system

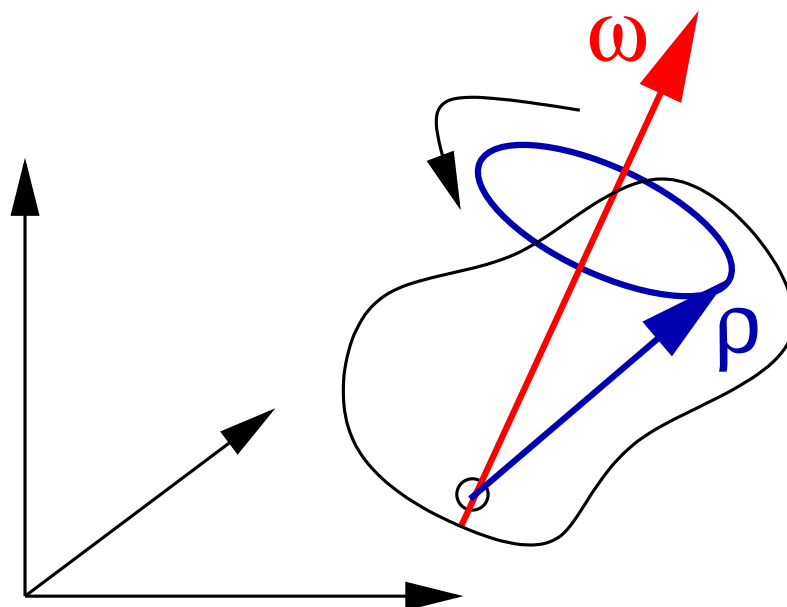
$$\frac{d\mathbf{r}}{dt} = \dot{\mathbf{R}}\boldsymbol{\rho} + \mathbf{R}\dot{\boldsymbol{\rho}}$$

but the second term is zero since we assumed $\boldsymbol{\rho}$ to be constant so we have

$$\frac{d\mathbf{r}}{dt} = \dot{\mathbf{R}}\mathbf{R}^T\mathbf{r}$$

Note that:

- $d\mathbf{r}/dt$ will have fixed magnitude;
- $d\mathbf{r}/dt$ will always be perpendicular to the axis of rotation;
- $d\mathbf{r}/dt$ will vary in direction within those constraints;
- $\mathbf{r}(t)$ will move in a plane in the fixed system.



Now let's consider the term $\dot{R}R^T$. First, note that $RR^T = I$ (the identity), so differentiating both sides yields

$$\begin{aligned}\dot{R}R^T + R\dot{R}^T &= 0 \\ \dot{R}R^T &= -R\dot{R}^T\end{aligned}$$

Thus $\dot{R}R^T$ is anti-symmetric:

$$\dot{R}R^T = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}$$

Now you can verify for yourself that application of a matrix of this form to an arbitrary vector has precisely the same effect as the cross product operator, $\boldsymbol{\omega} \times$, where $\boldsymbol{\omega} = [xyz]^T$. Loh-and-behold, we then we have

$$\dot{\mathbf{r}} = \boldsymbol{\omega} \times \mathbf{r}$$

matching the equation at the end of lecture 2, $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, as we would hope/expect.

3.5.1 Rotation: Part 2

Now suppose $\boldsymbol{\rho}$ is the position vector of a point P which **moves** in the rotating frame. There will be two contributions to motion with respect to the fixed frame, one due to its motion within the rotating frame, and one due to the rotation itself. So, returning to the equations we derived earlier:

$$\mathbf{r}(t) = R(t)\boldsymbol{\rho}(t)$$

and the instantaneous differential with respect to time:

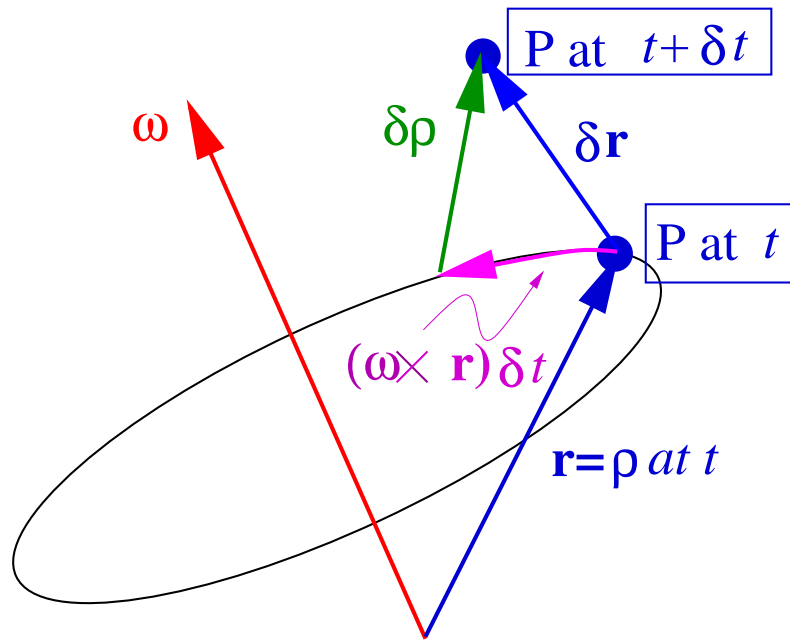
$$\frac{d\mathbf{r}}{dt} = \dot{R}\boldsymbol{\rho} + R\dot{\boldsymbol{\rho}} = \dot{R}R^T\mathbf{r} + R\dot{\boldsymbol{\rho}}$$

Now $\boldsymbol{\rho}$ is not constant, so its differential is not zero; hence rewriting this last equations we have that

The **instantaneous velocity** of P in the fixed frame is

$$\frac{d\mathbf{r}}{dt} = R\dot{\boldsymbol{\rho}} + \boldsymbol{\omega} \times \mathbf{r}$$

The second term of course, is the contribution from the rotating frame which we saw previously. The first is the linear velocity measured in the rotating frame $\dot{\boldsymbol{\rho}}$, referred to the fixed frame (via the rotation matrix R which aligns the two frames)



3.5.2 Rotation 3: Instantaneous acceleration

Our previous result is a general one relating the time derivatives of any vector in rotating and non-rotating frames. Let us now consider the second differential:

$$\ddot{\mathbf{r}} = \dot{\boldsymbol{\omega}} \times \mathbf{r} + \boldsymbol{\omega} \times \dot{\mathbf{r}} + \dot{R}\dot{\boldsymbol{\rho}} + R\ddot{\boldsymbol{\rho}}$$

We shall assume that the angular acceleration is zero, which kills off the first term, and so now, substituting for $\dot{\mathbf{r}}$ we have

$$\begin{aligned} \ddot{\mathbf{r}} &= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r} + R\dot{\boldsymbol{\rho}}) + \dot{R}\dot{\boldsymbol{\rho}} + R\ddot{\boldsymbol{\rho}} \\ &= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + \boldsymbol{\omega} \times R\dot{\boldsymbol{\rho}} + \dot{R}\dot{\boldsymbol{\rho}} + R\ddot{\boldsymbol{\rho}} \\ &= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + \boldsymbol{\omega} \times R\dot{\boldsymbol{\rho}} + \dot{R}(R^T R)\dot{\boldsymbol{\rho}} + R\ddot{\boldsymbol{\rho}} \\ &= \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + 2\boldsymbol{\omega} \times (R\dot{\boldsymbol{\rho}}) + R\ddot{\boldsymbol{\rho}} \end{aligned}$$

The **instantaneous acceleration** is therefore

$$\ddot{\mathbf{r}} = R\ddot{\boldsymbol{\rho}} + 2\boldsymbol{\omega} \times (R\dot{\boldsymbol{\rho}}) + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$$

- The first term is the acceleration of the point P in the rotating frame measured in the rotating frame, but referred to the fixed frame by the rotation R
- The last term is the centripetal acceleration due to the rotation. (Yes! Its magnitude is $\omega^2 r$ and its direction is that of $-\mathbf{r}$. Check it out.)

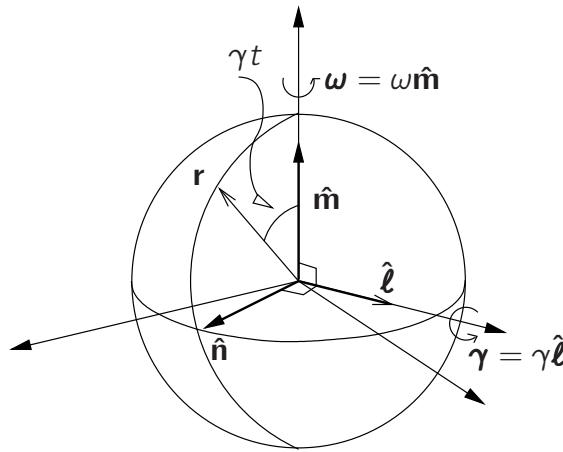


Figure 3.2: Coriolis example.

- The middle term is an extra term which arises because of the velocity of P in the rotating frame. It is known as the **Coriolis acceleration**, named after the French engineer who first identified it.

Because of the rotation of the earth, the Coriolis acceleration is of great importance in meteorology and accounts for the occurrence of high pressure anti-cyclones and low pressure cyclones in the northern hemisphere, in which the Coriolis acceleration is produced by a pressure gradient. It is also a very important component of the acceleration (hence the force exerted) by a rapidly moving robot arm, whose links whirl rapidly about rotary joints.

♣ Example

Q Find the instantaneous acceleration of a projectile fired along a line of longitude (with angular velocity of γ constant relative to the sphere) if the sphere is rotating with angular velocity ω .

A Consider a coordinate frame defined by mutually orthogonal unit vectors, \hat{l} , \hat{m} and \hat{n} , as shown in Fig. 3.2. We shall assume, without loss of generality, that the fixed and rotating frames are instantaneously aligned at the moment shown in the diagram, so that $R = I$, the identity, and hence $\mathbf{r} = \boldsymbol{\rho}$.

In the rotating frame

$$\dot{\boldsymbol{\rho}} = \boldsymbol{\gamma} \times \boldsymbol{\rho} \quad \text{and} \quad \ddot{\boldsymbol{\rho}} = \boldsymbol{\gamma} \times \dot{\boldsymbol{\rho}} = \boldsymbol{\gamma} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho})$$

So the in the fixed reference frame, because these two frames are instantaneously aligned

$$\ddot{\mathbf{r}} = \boldsymbol{\gamma} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) + 2\boldsymbol{\omega} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) .$$

The first term is the centripetal acceleration due to the projectile moving around the sphere — which it does because of the gravitational force. The

3.5. ROTATING SYSTEMS

last term is the centripetal acceleration resulting from the rotation of the sphere. The middle term is the Coriolis acceleration.

Using Fig. 3.2, at some instant t

$$\mathbf{r}(t) = \boldsymbol{\rho}(t) = r \cos(\gamma t) \hat{\mathbf{m}} + r \sin(\gamma t) \hat{\mathbf{n}}$$

and

$$\boldsymbol{\gamma} = \gamma \hat{\boldsymbol{\ell}}$$

Then

$$\boldsymbol{\gamma} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) = (\boldsymbol{\gamma} \cdot \boldsymbol{\rho}) \boldsymbol{\gamma} - \gamma^2 \boldsymbol{\rho} = -\gamma^2 \boldsymbol{\rho} = -\gamma^2 \mathbf{r},$$

Check the direction — the negative sign means it points *towards* the centre of the sphere, which is as expected.

Likewise the last term can be obtained as

$$\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = -\omega^2 r \sin(\gamma t) \hat{\mathbf{n}}$$

Note that it is perpendicular to the axis of rotation $\hat{\mathbf{m}}$, and because of the minus sign, directed towards the axis)

The Coriolis term is derived as:

$$\begin{aligned} 2\boldsymbol{\omega} \times \dot{\boldsymbol{\rho}} &= 2\boldsymbol{\omega} \times (\boldsymbol{\gamma} \times \boldsymbol{\rho}) \\ &= 2 \begin{bmatrix} 0 \\ \omega \\ 0 \end{bmatrix} \times \left(\begin{bmatrix} \gamma \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ r \cos \gamma t \\ r \sin \gamma t \end{bmatrix} \right) \\ &= 2\omega \gamma r \cos \gamma t \hat{\boldsymbol{\ell}} \end{aligned}$$

Instead of a projectile, now consider a rocket on rails which stretch north from the equator. As the rocket travels north it experiences the Coriolis force (exerted by the rails):

$$\begin{array}{ccccccc} 2 & \gamma & \omega & R \cos \gamma t & \hat{\boldsymbol{\ell}} \\ +ve & -ve & +ve & +ve & \end{array}$$

Hence the coriolis force is in the direction opposed to $\hat{\boldsymbol{\ell}}$ (i.e. in the opposite direction to the earth's rotation). In the absence of the rails (or atmosphere) the rocket's tangential speed (relative to the surface of the earth) is *greater* than the speed of the surface of the earth underneath it (since the radius of successive lines of latitude decreases) so it would (to an observer on the earth) appear to deflect to the east. The rails provide a coriolis force keeping it on the same meridian.

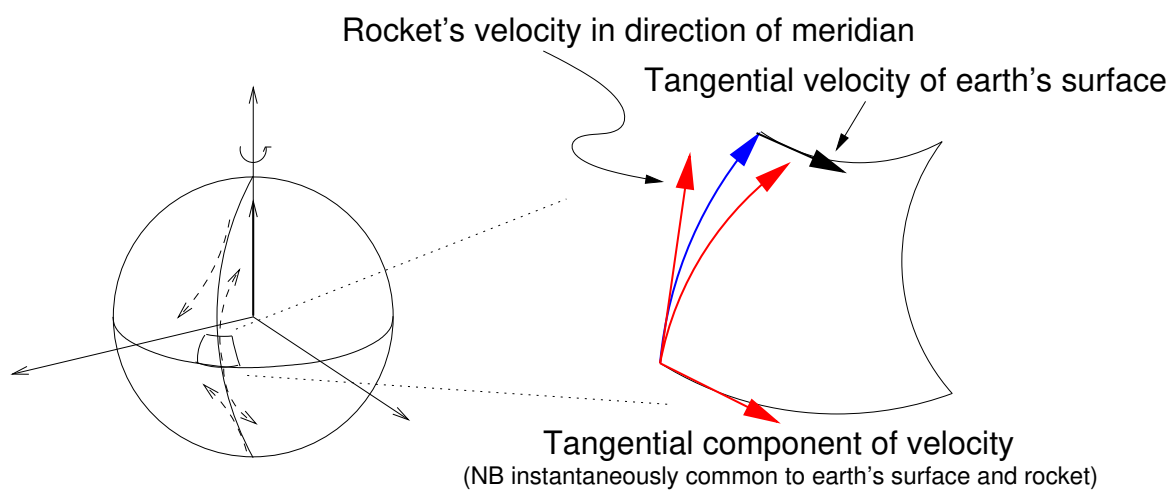


Figure 3.3: Rocket example

3.5. ROTATING SYSTEMS



Figure 3.4: Coriolis effect giving rise to weather systems

Lecture 4

Line, Surface and Volume Integrals. Curvilinear coordinates.

We started off the course being concerned with individual vectors **a**, **b**, **c**, and so on.

We went on to consider how single vectors vary over time or over some other parameter such as arc length.

In much of the rest of the course, we will be concerned with scalars and vectors which are defined over regions in space — scalar and vector *fields*

In this lecture we introduce line, surface and volume integrals, and consider how these are defined in non-Cartesian, *curvilinear coordinates*

4.1 Scalar and vector fields

When a scalar function $u(\mathbf{r})$ is determined or defined at each position \mathbf{r} in some region, we say that u is a **scalar field** in that region.

Similarly, if a vector function $\mathbf{v}(\mathbf{r})$ is defined at each point, then \mathbf{v} is a **vector field** in that region. As you will see, in **field theory** our aim is to derive statements about the bulk properties of scalar and vector fields, rather than to deal with individual scalars or vectors.

Familiar examples of each are shown in figure 4.1.

In Lecture 1 we worked out the force $\mathbf{F}(\mathbf{r})$ on a charge Q arising from a number of charges q_i . The electric field is \mathbf{F}/Q , so

$$\mathbf{E}(\mathbf{r}) = \sum_{i=1}^N K \frac{q_i}{|\mathbf{r} - \mathbf{r}_i|^3} (\mathbf{r} - \mathbf{r}_i) \quad . \quad \left(K = \frac{1}{4\pi\epsilon_r\epsilon_0} \right)$$

For example; you could work out the velocity field, in plane polars, at any point on



Figure 4.1: Examples of (a) a scalar field (pressure); (b) a vector field (wind velocity)

a wheel spinning about its axis

$$\mathbf{v}(\mathbf{r}) = \boldsymbol{\omega} \times \mathbf{r}$$

or the fluid flow field around a wing.

If the fields are independent of time, they are said to be *steady*. Of course, most vector fields of practical interest in engineering science are not steady, and some are unpredictable.

Let us first consider how to perform a variety of types of integration in vector and scalar fields.

4.2 Line integrals through fields

Line integrals are concerned with measuring the integrated interaction with a field as you move through it on some defined path. Eg, given a map showing the pollution density field in Oxford, you may wish to work out how much pollution you breathe in when cycling from college to the Department via different routes.

First recall the definition of an integral for a scalar function $f(x)$ of a single scalar variable x . One assumes a set of n samples $f_i = f(x_i)$ spaced by δx_i . One forms the limit of the sum of the products $f(x_i)\delta x_i$ as the number of samples tends to infinity

$$\int f(x)dx = \lim_{\substack{n \rightarrow \infty \\ \delta x_i \rightarrow 0}} \sum_{i=1}^n f_i \delta x_i .$$

For a smooth function, it is irrelevant how the function is subdivided.

4.2.1 Vector line integrals

In a vector line integral, the path L along which the integral is to be evaluated is split into a large number of *vector* segments $\delta \mathbf{r}_j$. Each line segment is then

4.2. LINE INTEGRALS THROUGH FIELDS

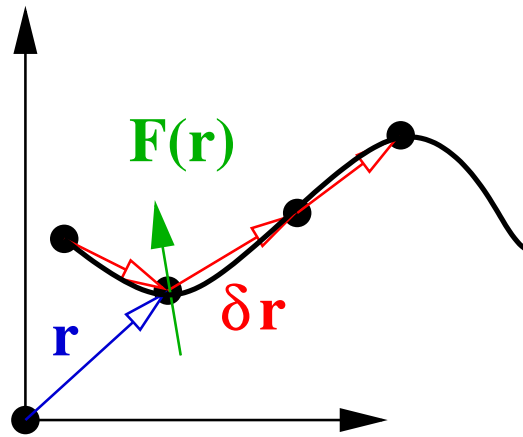


Figure 4.2: Line integral. In the diagram $\mathbf{F}(\mathbf{r})$ is a vector field, but it could be replaced with scalar field $U(\mathbf{r})$.

multiplied by the quantity associated with that point in space, the products are then summed and the limit taken as the lengths of the segments tend to zero.

There are three types of integral we have to think about, depending on the nature of the product:

1. Integrand $U(\mathbf{r})$ is a scalar field, hence the integral is a vector.

$$\mathbf{I} = \int_L U(\mathbf{r}) d\mathbf{r} \quad \left(= \lim_{\delta \mathbf{r}_i \rightarrow 0} \sum_i U_i \delta \mathbf{r}_i \right)$$

2. Integrand $\mathbf{a}(\mathbf{r})$ is a vector field dotted with $d\mathbf{r}$ hence the integral is a scalar:

$$I = \int_L \mathbf{a}(\mathbf{r}) \cdot d\mathbf{r} \quad \left(= \lim_{\delta \mathbf{r}_i \rightarrow 0} \sum_i \mathbf{a}_i \cdot \delta \mathbf{r}_i \right)$$

3. Integrand $\mathbf{a}(\mathbf{r})$ is a vector field crossed with $d\mathbf{r}$ hence vector result.

$$\mathbf{I} = \int_L \mathbf{a}(\mathbf{r}) \times d\mathbf{r} \quad \left(= \lim_{\delta \mathbf{r}_i \rightarrow 0} \sum_i \mathbf{a}_i \times \delta \mathbf{r}_i \right)$$

Note immediately that unlike an integral in a single scalar variable, there are many paths L from start point \mathbf{r}_A to end point \mathbf{r}_B , and the integral will in general depend on the path taken.

Physical examples of line integrals

- The total work done by a force \mathbf{F} as it moves a point from A to B along a given path C is given by a line integral of type 2 above. If the force acts

at point \mathbf{r} and the instantaneous displacement along curve C is $d\mathbf{r}$ then the infinitesimal work done is $dW = \mathbf{F} \cdot d\mathbf{r}$, and so the total work done traversing the path is

$$W_C = \int_C \mathbf{F} \cdot d\mathbf{r}$$

- Ampère's law relating magnetic field \mathbf{B} to linked current can be written as

$$\oint_C \mathbf{B} \cdot d\mathbf{r} = \mu_0 I$$

where I is the current enclosed by (closed) path C .

- The force on an element of wire carrying current I , placed in a magnetic field of strength \mathbf{B} , is $d\mathbf{F} = I d\mathbf{r} \times \mathbf{B}$. So if a loop this wire C is placed in the field then the total force will be and integral of type 3 above:

$$\mathbf{F} = I \oint_C d\mathbf{r} \times \mathbf{B}$$

Note that the expressions above are beautifully compact in vector notation, and are all independent of coordinate system. Of course when evaluating them we need to choose a coordinate system: often this is the standard Cartesian coordinate system (as in the worked examples below), but need not be, as we shall see in section 4.6.

♣ Examples

Q1 An example in the xy -plane. A force $\mathbf{F} = x^2 y \hat{i} + xy^2 \hat{j}$ acts on a body as it moves between $(0, 0)$ and $(1, 1)$.

Determine the work done when the path is

1. along the line $y = x$.
2. along the curve $y = x^n, n > 0$.
3. along the x axis to the point $(1, 0)$ and then along the line $x = 1$.

A1 This is an example of the “type 2” line integral. In planar Cartesians, $d\mathbf{r} = \hat{i}dx + \hat{j}dy$. Then the work done is

$$\int_L \mathbf{F} \cdot d\mathbf{r} = \int_L (x^2 y dx + xy^2 dy) .$$

1. For the path $y = x$ we find that $dy = dx$. So it is easiest to convert all y references to x .

$$\int_{(0,0)}^{(1,1)} (x^2 y dx + xy^2 dy) = \int_{x=0}^{x=1} (x^2 x dx + x x^2 dx) = \int_{x=0}^{x=1} 2x^3 dx = [x^4/2]_{x=0}^{x=1} = 1/2 .$$

4.2. LINE INTEGRALS THROUGH FIELDS

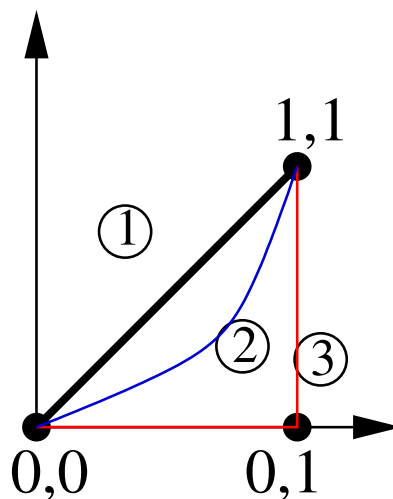


Figure 4.3: Line integral taken along three different paths

2. For the path $y = x^n$ we find that $dy = nx^{n-1}dx$, so again it is easiest to convert all y references to x .

$$\begin{aligned} \int_{(0,0)}^{(1,1)} (x^2y dx + xy^2 dy) &= \int_{x=0}^{x=1} (x^{n+2} dx + nx^{n-1} \cdot x \cdot x^{2n} dx) \\ &= \int_{x=0}^{x=1} (x^{n+2} dx + nx^{3n} dx) \\ &= \frac{1}{n+3} + \frac{n}{3n+1} \end{aligned}$$

3. This path is not smooth, so break it into two. Along the first section, $y = 0$ and $dy = 0$, and on the second $x = 1$ and $dx = 0$, so

$$\int_A^B (x^2y dx + xy^2 dy) = \int_{x=0}^{x=1} (x^2 \cdot 0 dx) + \int_{y=0}^{y=1} 1y^2 dy = 0 + [y^3/3]_{y=0}^{y=1} = 1/3.$$

So in general the integral depends on the path taken. Notice that answer (1) is the same as answer (2) when $n = 1$, and that answer (3) is the limiting value of answer (2) as $n \rightarrow \infty$.

Q2 Repeat part (2) using the Force $\mathbf{F} = xy^2\hat{i} + x^2y\hat{j}$.

A2 For the path $y = x^n$ we find that $dy = nx^{n-1}dx$, so

$$\begin{aligned} \int_{(0,0)}^{(1,1)} (y^2x dx + yx^2 dy) &= \int_{x=0}^{x=1} (x^{2n+1} dx + nx^{n-1} \cdot x^2 \cdot x^n dx) \\ &= \int_{x=0}^{x=1} (x^{2n+1} dx + nx^{2n+1} dx) \\ &= \frac{1}{2n+2} + \frac{n}{2n+2} \\ &= \frac{1}{2} \quad \text{independent of } n \end{aligned}$$

4.3 Line integrals in Conservative fields

In the second example, the line integral has the same value for the whole range of paths. In fact it is wholly independent of path. This is easy to see if we write $g(x, y) = x^2y^2/2$. Then using the definition of the perfect differential

$$dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy$$

we find that

$$\begin{aligned} \int_A^B (y^2x dx + yx^2 dy) &= \int_A^B dg \\ &= g_B - g_A \end{aligned}$$

which depends solely on the value of g at the start and end points, and not at all on the path used to get from A to B . Such a vector field is called **conservative**.

One sort of line integral performs the integration around a complete loop and is denoted with a ring. If \mathbf{E} is a conservative field, determine the value of

$$\oint \mathbf{E} \cdot d\mathbf{r}.$$

In electrostatics, if \mathbf{E} is the electric field then the potential function is

$$\phi = - \int \mathbf{E} \cdot d\mathbf{r}.$$

Do you think \mathbf{E} is conservative?

4.3.1 A note on line integrals defined in terms of arc length

Line integrals are often defined in terms of scalar arc length. They don't appear to involve vectors (but actually they are another form of type 2 defined earlier).

The integrals usually appear as follows

$$I = \int_L F(x, y, z) ds$$

and most often the path L is along a curve defined parametrically as $x = x(p)$, $y = y(p)$, $z = z(p)$ where p is some parameter. Convert the function to $F(p)$, writing

$$I = \int_{p_{\text{start}}}^{p_{\text{end}}} F(p) \frac{ds}{dp} dp$$

where

$$\frac{ds}{dp} = \left[\left(\frac{dx}{dp} \right)^2 + \left(\frac{dy}{dp} \right)^2 + \left(\frac{dz}{dp} \right)^2 \right]^{1/2}.$$

Note that the parameter p could be arc-length s itself, in which case $ds/dp = 1$ of course! Another possibility is that the parameter p is x — that is we are told $y = y(x)$ and $z = z(x)$. Then

$$I = \int_{x_{\text{start}}}^{x_{\text{end}}} F(x) \left[1 + \left(\frac{dy}{dx} \right)^2 + \left(\frac{dz}{dx} \right)^2 \right]^{1/2} dx.$$

4.4 Surface integrals

These can be defined by analogy with line integrals.

The surface S over which the integral is to be evaluated is now divided into infinitesimal vector elements of area $d\mathbf{S}$, the direction of the vector $d\mathbf{S}$ representing the direction of the surface normal and its magnitude representing the area of the element.

Again there are three possibilities:

- $\int_S U d\mathbf{S}$ — scalar field U ; vector integral.
- $\int_S \mathbf{a} \cdot d\mathbf{S}$ — vector field \mathbf{a} ; scalar integral.
- $\int_S \mathbf{a} \times d\mathbf{S}$ — vector field \mathbf{a} ; vector integral.

(in addition, of course, to the purely scalar form, $\int_S U dS$).

Physical example of surface integral

- Physical examples of surface integrals with vectors often involve the idea of *flux* of a vector field through a surface, $\int_S \mathbf{a} \cdot d\mathbf{S}$. For example the mass of fluid crossing a surface S in time dt is $dM = \rho \mathbf{v} \cdot d\mathbf{S} dt$ where $\rho(\mathbf{r})$ is the fluid density and $\mathbf{v}(\mathbf{r})$ is the fluid velocity. The total mass flux can be expressed as a surface integral:

$$\Phi_M = \int_S \rho(\mathbf{r}) \mathbf{v}(\mathbf{r}) \cdot d\mathbf{S}$$

Again, though this expression is coordinate free, we evaluate an example below using Cartesians. Note, however, that in some problems, symmetry may lead us to a different more natural coordinate system.

♣ Example

Evaluate $\int \mathbf{F} \cdot d\mathbf{S}$ over the $x = 1$ side of the cube shown in the figure when $\mathbf{F} = y\hat{\mathbf{i}} + z\hat{\mathbf{j}} + x\hat{\mathbf{k}}$.

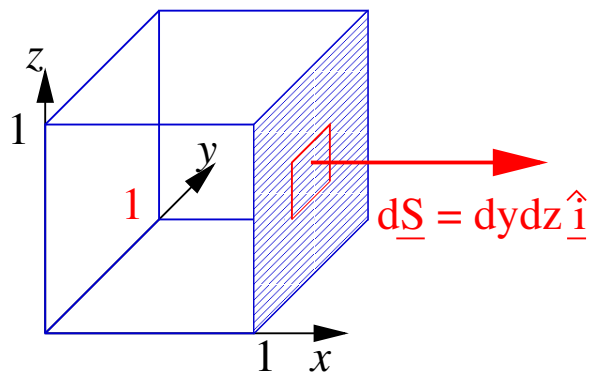
$d\mathbf{S}$ is perpendicular to the surface. Its \pm direction actually depends on the nature of the problem. More often than not, the surface will enclose a volume, and the surface direction is taken as everywhere emanating from the interior.

Hence for the $x = 1$ face of the cube

$$d\mathbf{S} = dydz\hat{\mathbf{i}}$$

and

$$\begin{aligned} \int \mathbf{F} \cdot d\mathbf{S} &= \int \int y dy dz \\ &= \frac{1}{2} y^2 \Big|_0^1 z \Big|_0^1 = \frac{1}{2} . \end{aligned}$$

**4.5 Volume integrals**

The definition of the volume integral is again taken as the limit of a sum of products as the size of the volume element tends to zero. One obvious difference though is that the element of volume is a scalar (how could you define a direction with an infinitesimal volume element?). The possibilities are:

4.6. CHANGING VARIABLES: CURVILINEAR COORDINATES

- $\int_V U(\mathbf{r})dV$ — scalar field; scalar integral.
- $\int_V \mathbf{a}dV$ — vector field; vector integral.

You have covered these (more or less) in your first year course, so not much more to say here. The next section considers these again in the context of a change of coordinates.

4.6 Changing variables: curvilinear coordinates

Up to now we have been concerned with Cartesian coordinates x, y, z with coordinate axes $\hat{i}, \hat{j}, \hat{k}$. When performing a line integral in Cartesian coordinates, you write

$$\mathbf{r} = x\hat{i} + y\hat{j} + z\hat{k} \quad \text{and} \quad d\mathbf{r} = dx\hat{i} + dy\hat{j} + dz\hat{k}$$

and can be sure that length scales are properly handled because – as we saw in Lecture 3 –

$$|d\mathbf{r}| = ds = \sqrt{dx^2 + dy^2 + dz^2}.$$

The reason for using the basis $\hat{i}, \hat{j}, \hat{k}$ rather than any other orthonormal basis set is that \hat{i} represents a direction in which x is increasing while the other two coordinates remain constant (and likewise for \hat{j} and \hat{k} with y and z respectively), simplifying the representation and resulting mathematics.

Often the symmetry of the problem strongly hints at using another coordinate system:

- likely to be plane, cylindrical, or spherical polars,
- but can be something more exotic

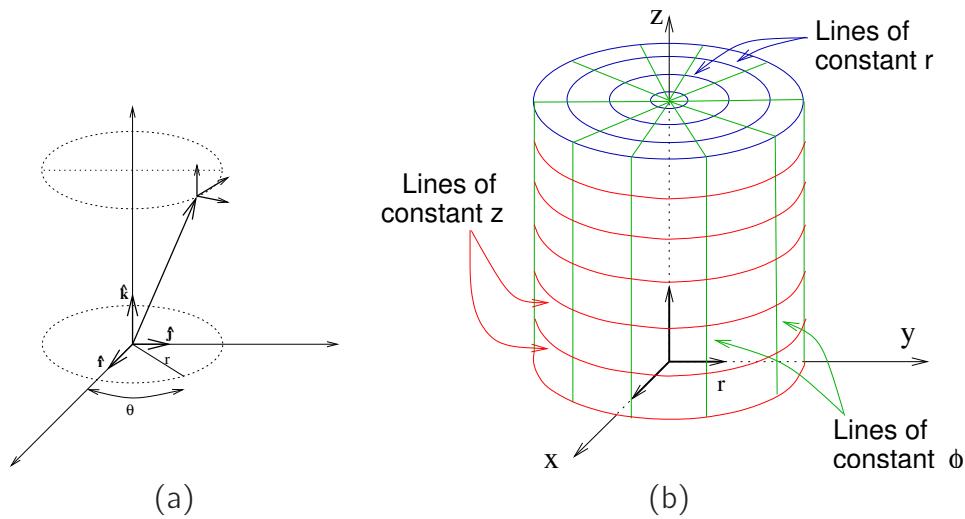
The general name for any different “ u, v, w ” coordinate system is a **curvilinear coordinate system**. We will see that the idea hinted at above – of defining a basis set by considering directions in which only one coordinate is (instantaneously) increasing – provides the appropriate generalisation.

We begin by discussing common special cases: cylindrical polars and spherical polars, and conclude with a more general formulation.

4.6.1 Cylindrical polar coordinates

As shown in figure 4.4 a point in space P having cartesian coordinates x, y, z can be expressed in terms of cylindrical polar coordinates, r, ϕ, z as follows:

$$\begin{aligned} \mathbf{r} &= x\hat{i} + y\hat{j} + z\hat{k} \\ &= r \cos \phi \hat{i} + r \sin \phi \hat{j} + z\hat{k} \end{aligned}$$

Figure 4.4: Cylindrical polars: (a) coordinate definition; (b) “iso” lines in r , ϕ and z .

Note that, by definition, $\frac{\partial \mathbf{r}}{\partial r}$ represents a direction in which (instantaneously) r is changing while the other two coordinates stay constant. That is, it is tangent to lines of constant ϕ and z . Likewise for $\frac{\partial \mathbf{r}}{\partial \phi}$ and $\frac{\partial \mathbf{r}}{\partial z}$. Thus the vectors:

$$\begin{aligned}\mathbf{e}_r &= \frac{\partial \mathbf{r}}{\partial r} = \cos \phi \hat{\mathbf{i}} + \sin \phi \hat{\mathbf{j}} \\ \mathbf{e}_\phi &= \frac{\partial \mathbf{r}}{\partial \phi} = -r \sin \phi \hat{\mathbf{i}} + r \cos \phi \hat{\mathbf{j}} \\ \mathbf{e}_z &= \frac{\partial \mathbf{r}}{\partial z} = \hat{\mathbf{k}}\end{aligned}$$

Aside on notation: some texts use the notation $\hat{\mathbf{r}}, \hat{\boldsymbol{\phi}}, \dots$ to represent the unit vectors that form the local basis set. Though I prefer the notation used here, where the basis vectors are written as $\hat{\mathbf{e}}$ with appropriate subscripts (as used in Riley *et al*), you should be aware of, and comfortable with, either possibility.

form a basis set in which we may describe infinitesimal vector displacements in the position of P , $d\mathbf{r}$. It is more usual, however, first to normalise the vectors to obtain their corresponding unit vectors, $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\phi$, $\hat{\mathbf{e}}_z$. Following the usual rules of calculus we may write:

$$\begin{aligned}d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial r} dr + \frac{\partial \mathbf{r}}{\partial \phi} d\phi + \frac{\partial \mathbf{r}}{\partial z} dz \\ &= dr \mathbf{e}_r + d\phi \mathbf{e}_\phi + dz \mathbf{e}_z \\ &= dr \hat{\mathbf{e}}_r + r d\phi \hat{\mathbf{e}}_\phi + dz \hat{\mathbf{e}}_z\end{aligned}$$

Now here is the important thing to note. In cartesian coordinates, a small change

4.6. CHANGING VARIABLES: CURVILINEAR COORDINATES

in (eg) x while keeping y and z constant would result in a displacement of

$$ds = |d\mathbf{r}| = \sqrt{d\mathbf{r} \cdot d\mathbf{r}} = \sqrt{dx^2 + 0 + 0} = dx$$

But in cylindrical polars, a small change in ϕ of $d\phi$ while keeping r and z constant results in a displacement of

$$ds = |d\mathbf{r}| = \sqrt{r^2(d\phi)^2} = r d\phi$$

Thus the size of the (infinitesimal) displacement is dependent on the value of r . Factors such as this r are known as **scale factors** or **metric coefficients**, and we must be careful to take them into account when, eg, performing line, surface or volume integrals, as you will below. For cylindrical polars the metric coefficients are clearly 1, r and 1.

Example: line integral in cylindrical coordinates

Q Evaluate $\oint_C \mathbf{a} \cdot d\mathbf{l}$, where $\mathbf{a} = x^3\hat{\mathbf{j}} - y^3\hat{\mathbf{i}} + x^2y\hat{\mathbf{k}}$ and C is the circle of radius r in the $z = 0$ plane, centred on the origin.

A Consider figure 4.5. In this case our cylindrical coordinates effectively reduce to plane polars since the path of integration is a circle in the $z = 0$ plane, but let's persist with the full set of coordinates anyway; the $\hat{\mathbf{k}}$ component of \mathbf{a} will play no role (it is normal to the path of integration and therefore cancels as seen below).

On the circle of interest

$$\mathbf{a} = r^3(-\sin^3\phi\hat{\mathbf{i}} + \cos^3\phi\hat{\mathbf{j}} + \cos^2\phi\sin\phi\hat{\mathbf{k}})$$

and (since $dz = dr = 0$ on the path)

$$\begin{aligned} d\mathbf{r} &= r d\phi \hat{\mathbf{e}}_\phi \\ &= r d\phi(-\sin\phi\hat{\mathbf{i}} + \cos\phi\hat{\mathbf{j}}) \end{aligned}$$

so that

$$\oint_C \mathbf{a} \cdot d\mathbf{r} = \int_0^{2\pi} r^4(\sin^4\phi + \cos^4\phi) d\phi = \frac{3\pi}{2} r^4$$

since

$$\int_0^{2\pi} \sin^4\phi d\phi = \int_0^{2\pi} \cos^4\phi d\phi = \frac{3\pi}{4}$$

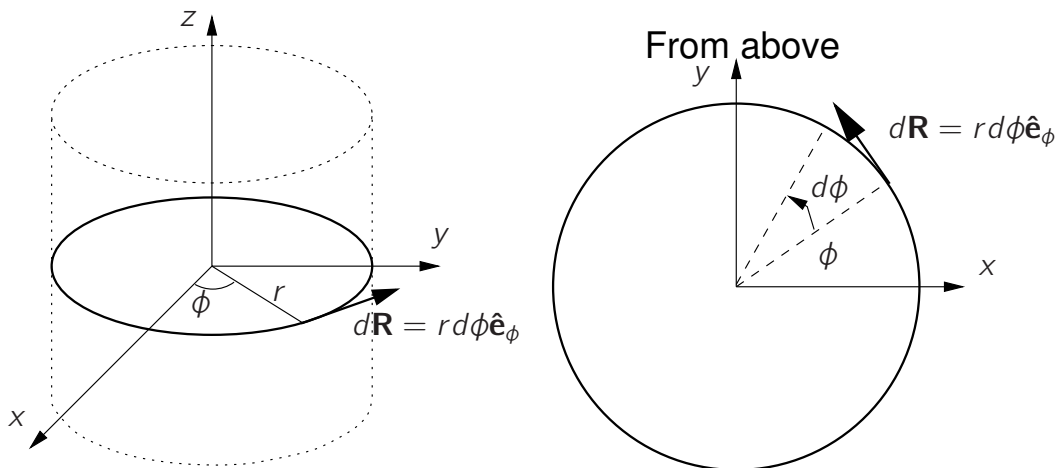


Figure 4.5: Line integral example in cylindrical coordinates

Volume integrals in cylindrical polars

In Cartesian coordinates a volume element is given by (see figure 4.6a):

$$dV = dx dy dz$$

Recall that the volume of a parallelopiped is given by the scalar triple product of the vectors which define it (see section 2.1.2). Thus the formula above can be derived (even though it is “obvious”) as:

$$dV = dx \hat{\mathbf{i}} \cdot (dy \hat{\mathbf{j}} \times dz \hat{\mathbf{k}}) = dx dy dz$$

since the basis set is orthonormal.

In cylindrical polars a volume element is given by (see figure 4.6b):

$$dV = dr \hat{\mathbf{e}}_r \cdot (r d\phi \hat{\mathbf{e}}_\phi \times dz \hat{\mathbf{e}}_z) = r d\phi dr dz$$

Note also that this volume, because it is a scalar triple product, can be written as a determinant:

$$dV = \begin{vmatrix} \hat{\mathbf{e}}_r dr \\ \hat{\mathbf{e}}_\phi r d\phi \\ \hat{\mathbf{e}}_z dz \end{vmatrix} = \begin{vmatrix} \mathbf{e}_r dr \\ \mathbf{e}_\phi d\phi \\ \mathbf{e}_z dz \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} & \frac{\partial z}{\partial \phi} \\ \frac{\partial x}{\partial z} & \frac{\partial y}{\partial z} & \frac{\partial z}{\partial z} \end{vmatrix} dr d\phi dz$$

where the equality on the right-hand side follows from the definitions of $\hat{\mathbf{e}}_r = \frac{\partial \mathbf{r}}{\partial r} = \frac{\partial x}{\partial r} \hat{\mathbf{i}} + \frac{\partial y}{\partial r} \hat{\mathbf{j}} + \frac{\partial z}{\partial r} \hat{\mathbf{k}}$, etc. This is the explanation for the “magical” appearance of the determinant in change-of-variables integration that you encountered in your first year maths!

4.6. CHANGING VARIABLES: CURVILINEAR COORDINATES

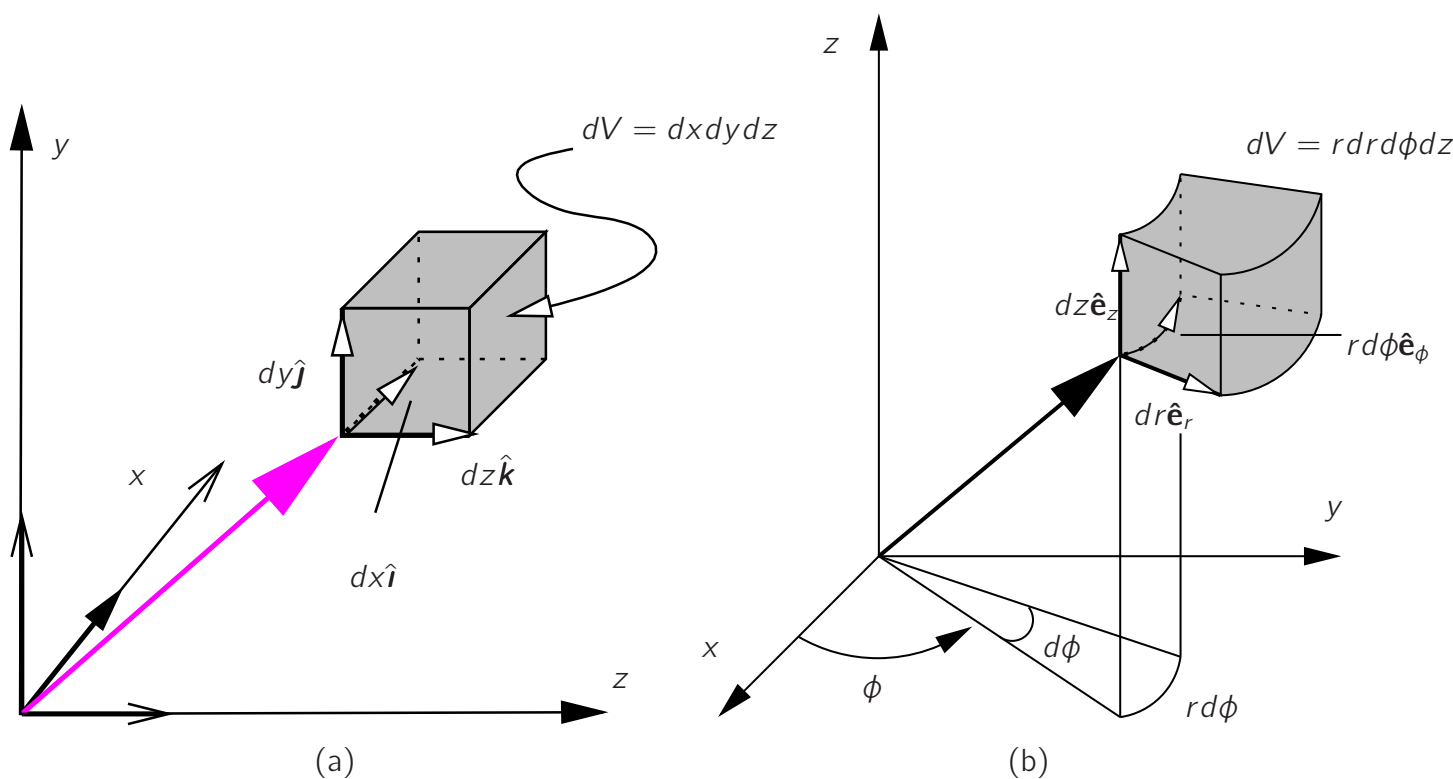


Figure 4.6: Volume elements dV in (a) Cartesian coordinates; (b) Cylindrical polar coordinates

Surface integrals in cylindrical polars

Recall from section 4.4 that for a surface element with normal along \hat{i} we have:

$$d\mathbf{S} = dydz\hat{i}$$

More explicitly this comes from finding normal to the plane that is tangent to the surface of constant x and from finding the area of an infinitesimal area element on the plane. In this case the plane is spanned by the vectors \hat{j} and \hat{k} and the area of the element given by (see section 1.3):

$$dS = |dy\hat{j} \times dz\hat{k}|$$

Thus

$$d\mathbf{S} = dy\hat{j} \times dz\hat{k} = \hat{i}dS = dydz\hat{i}$$

In cylindrical polars, surface area elements (see figure 4.7) are given by:

$$\begin{aligned} d\mathbf{S} &= dr\hat{e}_r \times r d\phi\hat{e}_\phi = r dr d\phi\hat{e}_z & (\text{for surfaces of constant } z) \\ d\mathbf{S} &= r d\phi\hat{e}_\phi \times dz\hat{e}_z = r d\phi dz\hat{e}_r & (\text{for surfaces of constant } r) \end{aligned}$$

Similarly we can find $d\mathbf{S}$ for surfaces of constant ϕ , though since these aren't as common this is left as a (relatively easy) exercise.

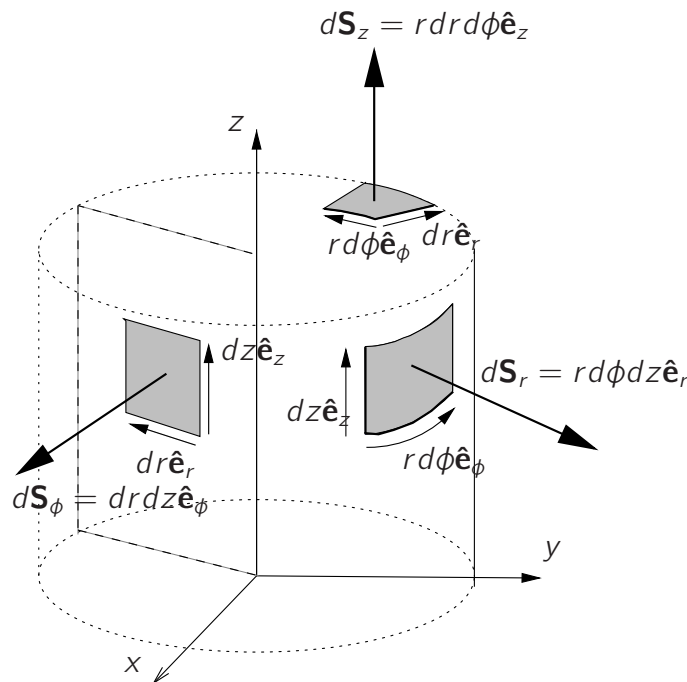


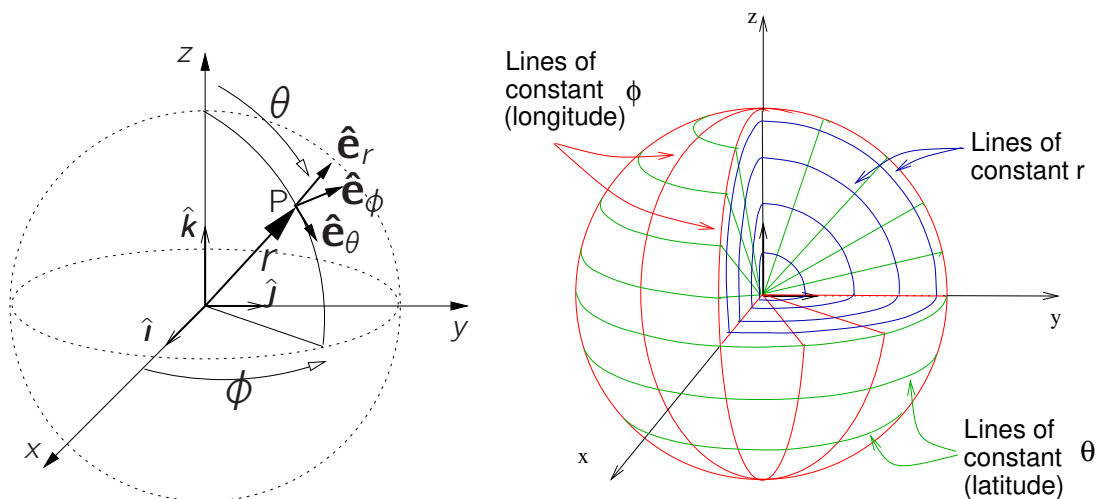
Figure 4.7: Surface elements in cylindrical polar coordinates

4.6.2 Spherical polars

Much of the development for spherical polars is similar to that for cylindrical polars. As shown in figure 4.6.2 a point in space P having cartesian coordinates x, y, z can be expressed in terms of spherical polar coordinates, r, θ, ϕ as follows:

$$\begin{aligned}\mathbf{r} &= x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} \\ &= r \sin \theta \cos \phi \hat{\mathbf{i}} + r \sin \theta \sin \phi \hat{\mathbf{j}} + r \cos \theta \hat{\mathbf{k}}\end{aligned}$$

The basis set in spherical polars is obtained in an analogous fashion: we find unit



4.6. CHANGING VARIABLES: CURVILINEAR COORDINATES

vectors which are in the direction of increase of each coordinate:

$$\begin{aligned}\mathbf{e}_r &= \frac{\partial \mathbf{r}}{\partial r} = \sin \theta \cos \phi \hat{\mathbf{i}} + \sin \theta \sin \phi \hat{\mathbf{j}} + \cos \theta \hat{\mathbf{k}} = \hat{\mathbf{e}}_r \\ \mathbf{e}_\theta &= \frac{\partial \mathbf{r}}{\partial \theta} = r \cos \theta \cos \phi \hat{\mathbf{i}} + r \cos \theta \sin \phi \hat{\mathbf{j}} - r \sin \theta \hat{\mathbf{k}} = r \hat{\mathbf{e}}_\theta \\ \mathbf{e}_\phi &= \frac{\partial \mathbf{r}}{\partial \phi} = -r \sin \theta \sin \phi \hat{\mathbf{i}} + r \sin \theta \cos \phi \hat{\mathbf{j}} = r \sin \theta \hat{\mathbf{e}}_\phi\end{aligned}$$

As with cylindrical polars, it is easily verified that the vectors $\hat{\mathbf{e}}_r, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_\phi$ form an orthonormal basis.

A small displacement $d\mathbf{r}$ is given by:

$$\begin{aligned}d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial r} dr + \frac{\partial \mathbf{r}}{\partial \theta} d\theta + \frac{\partial \mathbf{r}}{\partial \phi} d\phi \\ &= dr \mathbf{e}_r + d\theta \mathbf{e}_\theta + d\phi \mathbf{e}_\phi \\ &= dr \hat{\mathbf{e}}_r + r d\theta \hat{\mathbf{e}}_\theta + r \sin \theta d\phi \hat{\mathbf{e}}_\phi\end{aligned}$$

Thus the metric coefficients are $1, r, r \sin \theta$.

Volume integrals in spherical polars

In spherical polars a volume element is given by (see figure 4.8):

$$dV = dr \hat{\mathbf{e}}_r (r d\theta \hat{\mathbf{e}}_\theta \times r \sin \theta d\phi \hat{\mathbf{e}}_\phi) = r^2 \sin \theta dr d\theta d\phi$$

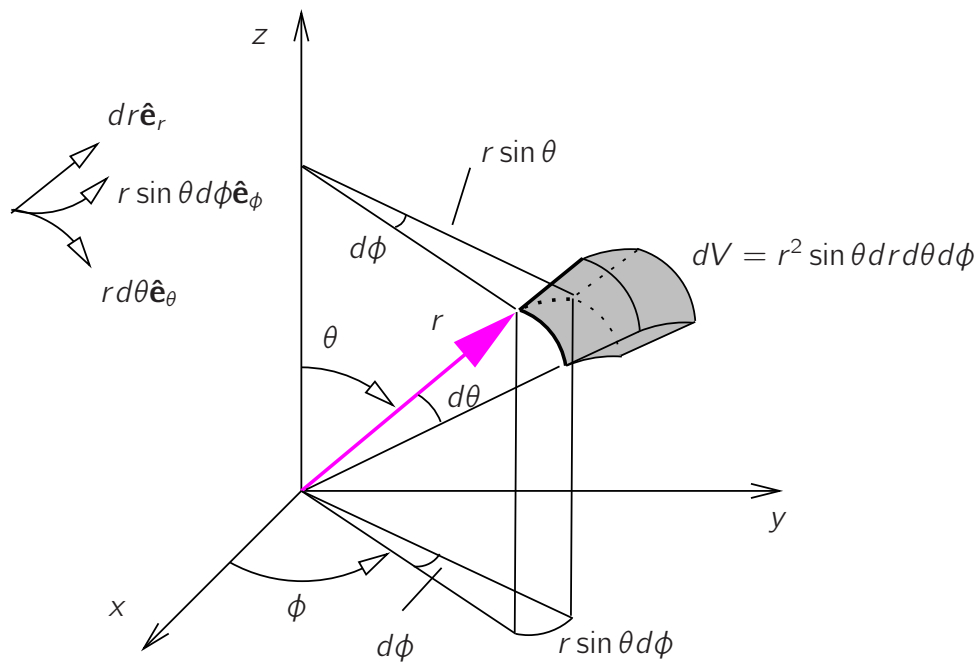
Note again that this volume could be written as a determinant, but this is left as an exercise.

Surface integrals in spherical polars

The most (the only?) useful surface elements in spherical polars are those tangent to surfaces of constant r (see figure 4.9). The surface direction (unnormalised) is given by $\hat{\mathbf{e}}_\theta \times \hat{\mathbf{e}}_\phi = \hat{\mathbf{e}}_r$ and the area of an infinitesimal surface element is given by $|r d\theta \hat{\mathbf{e}}_\theta \times r \sin \theta d\phi \hat{\mathbf{e}}_\phi| = r^2 \sin \theta d\theta d\phi$.

Thus a surface element $d\mathbf{S}$ in spherical polars is given by

$$d\mathbf{S} = r d\theta \hat{\mathbf{e}}_\theta \times r \sin \theta d\phi \hat{\mathbf{e}}_\phi = r^2 \sin \theta \hat{\mathbf{e}}_r$$

Figure 4.8: Volume element dV in spherical polar coordinates♣ **Example: surface integral in spherical polars**

Q Evaluate $\int_S \mathbf{a} \cdot d\mathbf{S}$, where $\mathbf{a} = z^3 \hat{\mathbf{k}}$ and S is the sphere of radius A centred on the origin.

A On the surface of the sphere:

$$\mathbf{a} = A^3 \cos^3 \theta \hat{\mathbf{k}} \quad d\mathbf{S} = A^2 \sin \theta \, d\theta \, d\phi \hat{\mathbf{e}}_r$$

Hence

$$\begin{aligned} \int_S \mathbf{a} \cdot d\mathbf{S} &= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} A^3 \cos^3 \theta \, A^2 \sin \theta \, [\hat{\mathbf{e}}_r \cdot \hat{\mathbf{k}}] \, d\theta \, d\phi \\ &= A^5 \int_0^{2\pi} d\phi \int_0^{\pi} \cos^3 \theta \sin \theta [\cos \theta] \, d\theta \\ &= 2\pi A^5 \frac{1}{5} [-\cos^5 \theta]_0^{\pi} \\ &= \frac{4\pi A^5}{5} \end{aligned}$$

4.6. CHANGING VARIABLES: CURVILINEAR COORDINATES

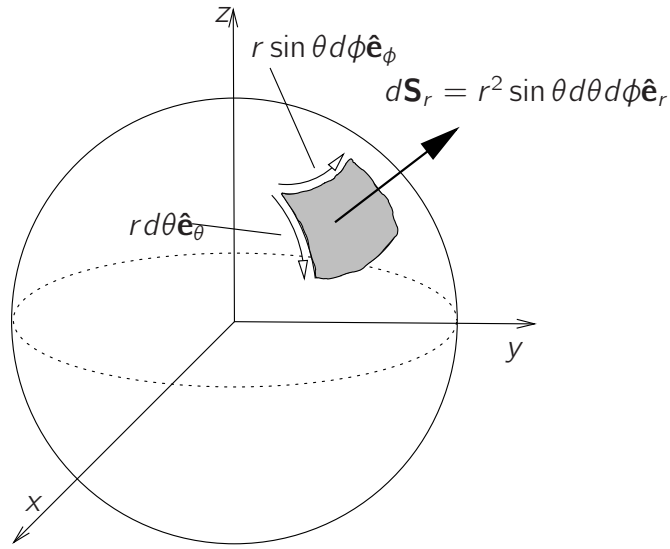


Figure 4.9: Surface element $d\mathbf{S}$ in spherical polar coordinates

4.6.3 General curvilinear coordinates

Cylindrical and spherical polar coordinates are two (useful) examples of general curvilinear coordinates. In general a point P with Cartesian coordinates x, y, z can be expressed in terms of the curvilinear coordinates u, v, w where

$$x = x(u, v, w), \quad y = y(u, v, w), \quad z = z(u, v, w)$$

Thus

$$\mathbf{r} = x(u, v, w)\hat{\mathbf{i}} + y(u, v, w)\hat{\mathbf{j}} + z(u, v, w)\hat{\mathbf{k}}$$

and

$$\frac{\partial \mathbf{r}}{\partial u} = \frac{\partial x}{\partial u}\hat{\mathbf{i}} + \frac{\partial y}{\partial u}\hat{\mathbf{j}} + \frac{\partial z}{\partial u}\hat{\mathbf{k}}$$

and similarly for partials with respect to v and w , so

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}du + \frac{\partial \mathbf{r}}{\partial v}dv + \frac{\partial \mathbf{r}}{\partial w}dw$$

We now define the local coordinate system as before by considering the directions in which each coordinate “unilaterally” (and instantaneously) increases:

$$\mathbf{e}_u = \frac{\partial \mathbf{r}}{\partial u} = \left| \frac{\partial \mathbf{r}}{\partial u} \right| \hat{\mathbf{e}}_u = h_u \hat{\mathbf{e}}_u$$

$$\mathbf{e}_v = \frac{\partial \mathbf{r}}{\partial v} = \left| \frac{\partial \mathbf{r}}{\partial v} \right| \hat{\mathbf{e}}_v = h_v \hat{\mathbf{e}}_v$$

$$\mathbf{e}_w = \frac{\partial \mathbf{r}}{\partial w} = \left| \frac{\partial \mathbf{r}}{\partial w} \right| \hat{\mathbf{e}}_w = h_w \hat{\mathbf{e}}_w$$

LECTURE 4. LINE, SURFACE AND VOLUME INTEGRALS. CURVILINEAR COORDINATES.

The metric coefficients are therefore $h_u = \left| \frac{\partial \mathbf{r}}{\partial u} \right|$, $h_v = \left| \frac{\partial \mathbf{r}}{\partial v} \right|$ and $h_w = \left| \frac{\partial \mathbf{r}}{\partial w} \right|$.

A volume element is in general given by

$$dV = h_u du \hat{\mathbf{e}}_u \cdot (h_v dv \hat{\mathbf{e}}_v \times h_w dw \hat{\mathbf{e}}_w)$$

and simplifies *if the coordinate system is orthonormal* (since $\hat{\mathbf{e}}_u \cdot (\hat{\mathbf{e}}_v \times \hat{\mathbf{e}}_w) = 1$) to

$$dV = h_u h_v h_w du dv dw$$

A surface element (normal to constant w , say) is in general

$$d\mathbf{S} = h_u du \hat{\mathbf{e}}_u \times h_v dv \hat{\mathbf{e}}_v$$

and simplifies *if the coordinate system is orthogonal* to

$$d\mathbf{S} = h_u h_v du dv \hat{\mathbf{e}}_w$$

4.6.4 Summary

To summarise:

General curvilinear coordinates

$$x = x(u, v, w), \quad y = y(u, v, w), \quad z = z(u, v, w)$$

$$\mathbf{r} = x(u, v, w)\hat{\mathbf{i}} + y(u, v, w)\hat{\mathbf{j}} + z(u, v, w)\hat{\mathbf{k}}$$

$$h_u = \left| \frac{\partial \mathbf{r}}{\partial u} \right|, \quad h_v = \left| \frac{\partial \mathbf{r}}{\partial v} \right|, \quad h_w = \left| \frac{\partial \mathbf{r}}{\partial w} \right|$$

$$\hat{\mathbf{u}} = \hat{\mathbf{e}}_u = \frac{1}{h_u} \frac{\partial \mathbf{r}}{\partial u}, \quad \hat{\mathbf{v}} = \hat{\mathbf{e}}_v = \frac{1}{h_v} \frac{\partial \mathbf{r}}{\partial v}, \quad \hat{\mathbf{w}} = \hat{\mathbf{e}}_w = \frac{1}{h_w} \frac{\partial \mathbf{r}}{\partial w}$$

$$d\mathbf{r} = h_u du \hat{\mathbf{u}} + h_v dv \hat{\mathbf{v}} + h_w dw \hat{\mathbf{w}}$$

$$dV = h_u h_v h_w du dv dw \hat{\mathbf{u}} \cdot (\hat{\mathbf{v}} \times \hat{\mathbf{w}})$$

$$d\mathbf{S} = h_u h_v du dv \hat{\mathbf{u}} \times \hat{\mathbf{v}} \quad (\text{for surface element tangent to constant } w)$$

Plane polar coordinates

$$x = r \cos \theta, \quad y = r \sin \theta$$

$$\mathbf{r} = r \cos \theta \hat{\mathbf{i}} + r \sin \theta \hat{\mathbf{j}}$$

$$h_r = 1, \quad h_\theta = r$$

$$\hat{\mathbf{e}}_r = \cos \theta \hat{\mathbf{i}} + \sin \theta \hat{\mathbf{j}}, \quad \hat{\mathbf{e}}_\theta = -\sin \theta \hat{\mathbf{i}} + \cos \theta \hat{\mathbf{j}}$$

$$d\mathbf{r} = dr \hat{\mathbf{e}}_r + r d\theta \hat{\mathbf{e}}_\theta$$

$$d\mathbf{S} = r dr d\theta \hat{\mathbf{k}}$$

4.6. CHANGING VARIABLES: CURVILINEAR COORDINATES

Cylindrical polar coordinates

$$\begin{aligned}
 x &= r \cos \phi, & y &= r \sin \phi, & z &= z \\
 \mathbf{r} &= r \cos \phi \hat{\mathbf{i}} + r \sin \phi \hat{\mathbf{j}} + z \hat{\mathbf{k}} \\
 h_r &= 1, & h_\phi &= r, & h_z &= 1 \\
 \hat{\mathbf{e}}_r &= \cos \phi \hat{\mathbf{i}} + \sin \phi \hat{\mathbf{j}}, & \hat{\mathbf{e}}_\phi &= -\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}}, & \hat{\mathbf{e}}_z &= \hat{\mathbf{k}} \\
 d\mathbf{r} &= dr \hat{\mathbf{e}}_r + r d\phi \hat{\mathbf{e}}_\phi + dz \hat{\mathbf{e}}_z \\
 d\mathbf{S} &= r dr d\phi \hat{\mathbf{k}} \quad (\text{on the flat ends}) \\
 d\mathbf{S} &= r d\phi dz \hat{\mathbf{e}}_r \quad (\text{on the curved sides}) \\
 dV &= r dr d\phi dz
 \end{aligned}$$

Spherical polar coordinates

$$\begin{aligned}
 x &= r \sin \theta \cos \phi, & y &= r \sin \theta \sin \phi, & z &= r \cos \theta \\
 \mathbf{r} &= r \sin \theta \cos \phi \hat{\mathbf{i}} + r \sin \theta \sin \phi \hat{\mathbf{j}} + r \cos \theta \hat{\mathbf{k}} \\
 h_r &= 1, & h_\theta &= r, & h_\phi &= r \sin \theta \\
 \hat{\mathbf{e}}_r &= \sin \theta \cos \phi \hat{\mathbf{i}} + \sin \theta \sin \phi \hat{\mathbf{j}} + \cos \theta \hat{\mathbf{k}} \\
 \hat{\mathbf{e}}_\theta &= \cos \theta \cos \phi \hat{\mathbf{i}} + \cos \theta \sin \phi \hat{\mathbf{j}} + \sin \theta \hat{\mathbf{k}} \\
 \hat{\mathbf{e}}_\phi &= -\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}} \\
 d\mathbf{r} &= dr \hat{\mathbf{e}}_r + r d\theta \hat{\mathbf{e}}_\theta + r \sin \theta d\phi \hat{\mathbf{e}}_\phi \\
 d\mathbf{S} &= r^2 \sin \theta dr d\theta d\phi \hat{\mathbf{e}}_r \quad (\text{on a spherical surface}) \\
 dV &= r^2 \sin \theta dr d\theta d\phi
 \end{aligned}$$

Lecture 5

Vector Operators: Grad, Div and Curl

In the first lecture of the second part of this course we move more to consider properties of fields. We introduce three field operators which reveal interesting collective field properties, viz.

- the **gradient** of a scalar field,
- the **divergence** of a vector field, and
- the **curl** of a vector field.

There are two points to get over about each:

- The mechanics of taking the grad, div or curl, for which you will need to brush up your multivariate calculus.
- The underlying physical meaning — that is, why they are worth bothering about.

In Lecture 6 we will look at combining these vector operators.

5.1 The gradient of a scalar field

Recall the discussion of temperature distribution throughout a room in the overview, where we wondered how a scalar would vary as we moved off in an arbitrary direction. Here we find out how.

If $U(x, y, z)$ is a scalar field, ie a scalar function of position $\mathbf{r} = [x, y, z]$ in 3 dimensions, then its **gradient** at any point is defined in Cartesian co-ordinates by

$$\text{grad}U = \frac{\partial U}{\partial x} \hat{i} + \frac{\partial U}{\partial y} \hat{j} + \frac{\partial U}{\partial z} \hat{k} .$$

It is usual to define the **vector operator** which is called “del” or “nabla”

$$\nabla = \hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} .$$

Then

$$\text{grad}U \equiv \nabla U .$$

Note immediately that ∇U is a vector field!

Without thinking too carefully about it, we can see that the gradient of a scalar field tends to point in the direction of greatest change of the field. Later we will be more precise.

♣ Worked examples of gradient evaluation

1. $U = x^2$

$$\Rightarrow \nabla U = \left(\frac{\partial}{\partial x} \hat{i} + \frac{\partial}{\partial y} \hat{j} + \frac{\partial}{\partial z} \hat{k} \right) x^2 = 2x \hat{i} .$$

2. $U = r^2$

$$\begin{aligned} r^2 &= x^2 + y^2 + z^2 \\ \Rightarrow \nabla U &= \left(\frac{\partial}{\partial x} \hat{i} + \frac{\partial}{\partial y} \hat{j} + \frac{\partial}{\partial z} \hat{k} \right) (x^2 + y^2 + z^2) \\ &= 2x \hat{i} + 2y \hat{j} + 2z \hat{k} = 2 \mathbf{r} . \end{aligned}$$

3. $U = \mathbf{c} \cdot \mathbf{r}$, where \mathbf{c} is constant.

$$\Rightarrow \nabla U = \left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) (c_1 x + c_2 y + c_3 z) = c_1 \hat{i} + c_2 \hat{j} + c_3 \hat{k} = \mathbf{c} .$$

4. $U = f(r)$, where $r = \sqrt{x^2 + y^2 + z^2}$

U is a function of r alone so df/dr exists. As $U = f(x, y, z)$ also,

$$\frac{\partial f}{\partial x} = \frac{df}{dr} \frac{\partial r}{\partial x} \quad \frac{\partial f}{\partial y} = \frac{df}{dr} \frac{\partial r}{\partial y} \quad \frac{\partial f}{\partial z} = \frac{df}{dr} \frac{\partial r}{\partial z} .$$

$$\Rightarrow \nabla U = \frac{\partial f}{\partial x} \hat{i} + \frac{\partial f}{\partial y} \hat{j} + \frac{\partial f}{\partial z} \hat{k} = \frac{df}{dr} \left(\frac{\partial r}{\partial x} \hat{i} + \frac{\partial r}{\partial y} \hat{j} + \frac{\partial r}{\partial z} \hat{k} \right)$$

But $r = \sqrt{x^2 + y^2 + z^2}$, so $\partial r / \partial x = x/r$ and similarly for y, z .

$$\Rightarrow \nabla U = \frac{df}{dr} \left(\frac{x \hat{i} + y \hat{j} + z \hat{k}}{r} \right) = \frac{df}{dr} \left(\frac{\mathbf{r}}{r} \right) .$$

5.2. THE SIGNIFICANCE OF GRAD

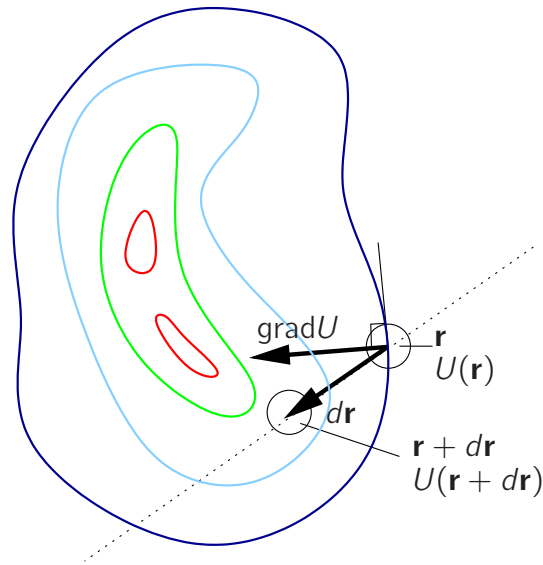


Figure 5.1: The directional derivative

5.2 The significance of grad

If our current position is \mathbf{r} in some scalar field U (Fig. 5.1), and we move an infinitesimal distance $d\mathbf{r}$, we know that the change in U is

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy + \frac{\partial U}{\partial z}dz .$$

But we know that $d\mathbf{r} = (\hat{i}dx + \hat{j}dy + \hat{k}dz)$ and $\nabla U = (\hat{i}\partial U/\partial x + \hat{j}\partial U/\partial y + \hat{k}\partial U/\partial z)$, so that the change in U is also given by the scalar product

$$dU = \nabla U \cdot d\mathbf{r} .$$

Now divide both sides by ds

$$\frac{dU}{ds} = \nabla U \cdot \frac{d\mathbf{r}}{ds} .$$

But remember that $|d\mathbf{r}| = ds$, so $d\mathbf{r}/ds$ is a unit vector in the direction of $d\mathbf{r}$.

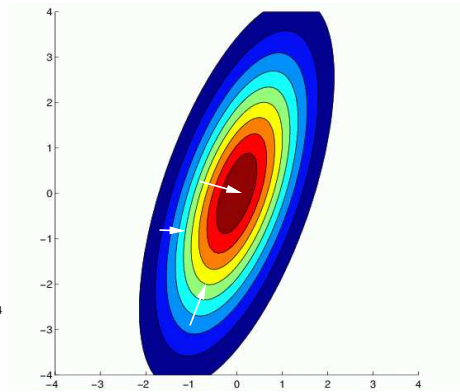
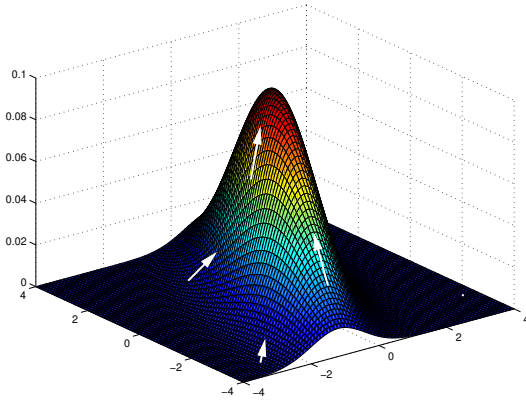
This result can be paraphrased as:

- $\text{grad}U$ has the property that the rate of change of U wrt distance in a particular direction ($\hat{\mathbf{d}}$) is the projection of $\text{grad}U$ onto that direction (or the component of $\text{grad}U$ in that direction).

The quantity dU/ds is called a **directional derivative**, but note that in general it has a different value for each direction, and so has no meaning until you specify the direction.

We could also say that

- At any point P , $\text{grad}U$ points in the direction of greatest change of U at P , and has magnitude equal to the rate of change of U wrt distance in that direction.



Another nice property emerges if we think of a surface of constant U – that is the locus (x, y, z) for

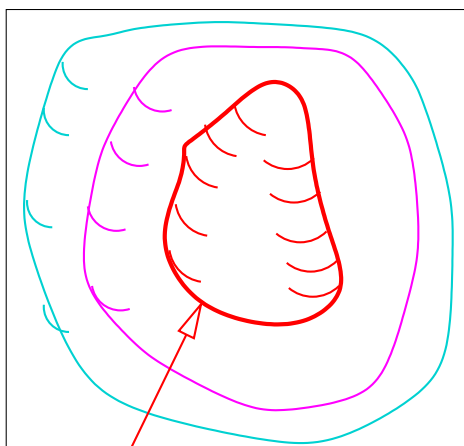
$$U(x, y, z) = \text{constant} .$$

If we move a tiny amount within that iso- U surface, there is no change in U , so $dU/ds = 0$. So for any $d\mathbf{r}/ds$ in the surface

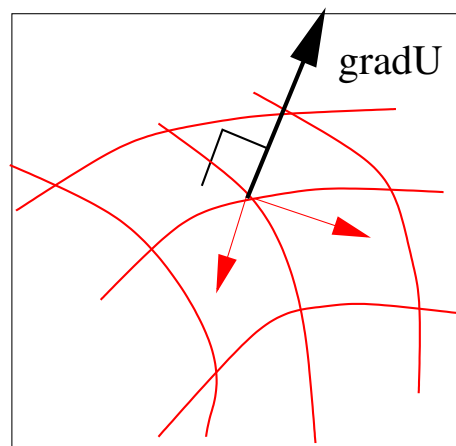
$$\nabla U \cdot \frac{d\mathbf{r}}{ds} = 0 .$$

But $d\mathbf{r}/ds$ is a tangent to the surface, so this result shows that

- $\text{grad}U$ is everywhere NORMAL to a surface of constant U .



Surface of constant U
These are called Level Surfaces



Surface of constant U

5.3 The divergence of a vector field

The divergence computes a scalar quantity from a vector field by differentiation.

If $\mathbf{a}(x, y, z)$ is a vector function of position in 3 dimensions, that is $\mathbf{a} = a_1\hat{i} + a_2\hat{j} + a_3\hat{k}$, then its divergence at any point is defined in Cartesian co-ordinates by

$$\text{div}\mathbf{a} = \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y} + \frac{\partial a_3}{\partial z}$$

We can write this in a simplified notation using a scalar product with the ∇ vector differential operator:

$$\text{div}\mathbf{a} = \left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) \cdot \mathbf{a} = \nabla \cdot \mathbf{a}$$

Notice that the divergence of a vector field is a scalar field.

♣ Examples of divergence evaluation

| | \mathbf{a} | $\text{div}\mathbf{a}$ |
|----|--|-----------------------------------|
| 1) | $x\hat{i}$ | 1 |
| 2) | $\mathbf{r}(= x\hat{i} + y\hat{j} + z\hat{k})$ | 3 |
| 3) | \mathbf{r}/r^3 | 0 |
| 4) | $r\mathbf{c}$, for \mathbf{c} constant | $(\mathbf{r} \cdot \mathbf{c})/r$ |

We work through example 3).

The x component of \mathbf{r}/r^3 is $x.(x^2 + y^2 + z^2)^{-3/2}$, and we need to find $\partial/\partial x$ of it.

$$\begin{aligned} \frac{\partial}{\partial x} x.(x^2 + y^2 + z^2)^{-3/2} &= 1.(x^2 + y^2 + z^2)^{-3/2} + x \frac{-3}{2} (x^2 + y^2 + z^2)^{-5/2} . 2x \\ &= r^{-3} (1 - 3x^2 r^{-2}) . \end{aligned}$$

The terms in y and z are similar, so that

$$\begin{aligned} \text{div}(\mathbf{r}/r^3) &= r^{-3} (3 - 3(x^2 + y^2 + z^2)r^{-2}) = r^{-3} (3 - 3) \\ &= 0 \end{aligned}$$

5.4 The significance of div

Consider a typical vector field, water flow, and denote it by $\mathbf{a}(\mathbf{r})$. This vector has magnitude equal to the mass of water crossing a unit area perpendicular to the direction of \mathbf{a} per unit time.

Now take an infinitesimal volume element dV and figure out the balance of the flow of \mathbf{a} in and out of dV .

To be specific, consider the volume element $dV = dx dy dz$ in Cartesian co-ordinates, and think first about the face of area $dx dz$ perpendicular to the y axis and facing outwards in the negative y direction. (That is, the one with surface area $d\mathbf{S} = -dx dz \hat{\mathbf{j}}$.)

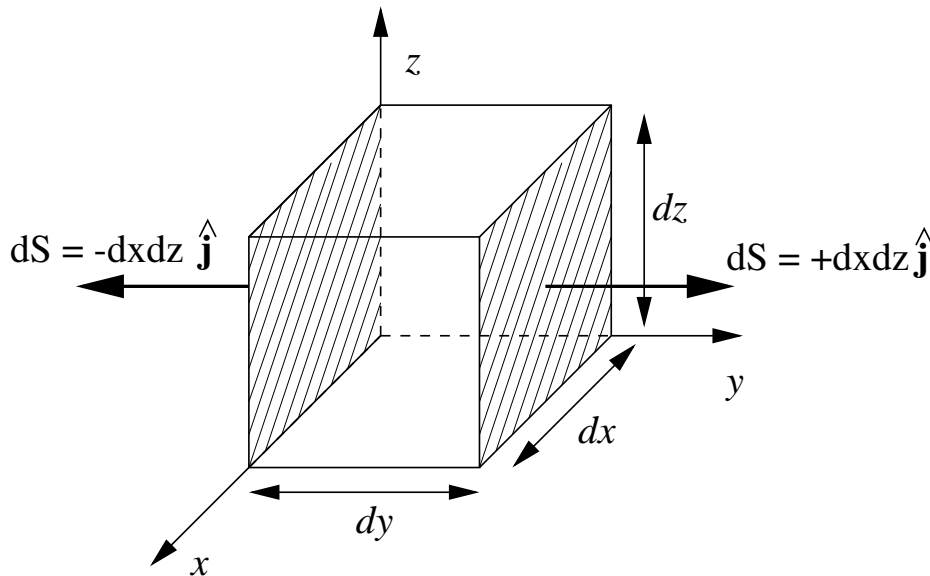


Figure 5.2: Elemental volume for calculating divergence.

The component of the vector \mathbf{a} normal to this face is $\mathbf{a} \cdot \hat{\mathbf{j}} = a_y$, and is pointing inwards, and so its contribution to the OUTWARD flux from this surface is

$$\mathbf{a} \cdot d\mathbf{S} = -a_y(y) dz dx ,$$

where $a_y(y)$ means that a_y is a function of y . (By the way, flux here denotes mass per unit time.)

A similar contribution, but of opposite sign, will arise from the opposite face, but we must remember that we have moved along y by an amount dy , so that this OUTWARD amount is

$$a_y(y + dy) dz dx = \left(a_y + \frac{\partial a_y}{\partial y} dy \right) dx dz$$

The total outward amount from these two faces is

$$\frac{\partial a_y}{\partial y} dy dx dz = \frac{\partial a_y}{\partial y} dV$$

Summing the other faces gives a total outward flux of

$$\left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) dV = \nabla \cdot \mathbf{a} dV$$

So we see that

5.5. THE LAPLACIAN: $\text{DIV}(\text{GRAD}U)$ OF A SCALAR FIELD

The divergence of a vector field represents the flux generation per unit volume at each point of the field. (Divergence because it is an efflux not an influx.)

Interestingly we also saw that the total efflux from the infinitesimal volume was equal to the flux integrated over the surface of the volume.

(NB: The above does not constitute a rigorous proof of the assertion because we have not proved that the quantity calculated is independent of the co-ordinate system used, but it will suffice for our purposes.)

5.5 The Laplacian: $\text{div}(\text{grad}U)$ of a scalar field

Recall that $\text{grad}U$ of *any* scalar field U is a vector field. Recall also that we can compute the divergence of any vector field. So we can certainly compute $\text{div}(\text{grad}U)$, even if we don't know what it means yet.

Here is where the ∇ operator starts to be really handy.

$$\begin{aligned}\nabla \cdot (\nabla U) &= \left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) \cdot \left(\left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) U \right) \\ &= \left(\left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) \cdot \left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) \right) U \\ &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) U \\ &= \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} \right)\end{aligned}$$

This last expression occurs frequently in engineering science (you will meet it next in solving Laplace's Equation in partial differential equations). For this reason, the operator ∇^2 is called the "Laplacian"

$$\nabla^2 U = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) U$$

Laplace's equation itself is

$$\nabla^2 U = 0$$

♣ Examples of $\nabla^2 U$ evaluation

| U | $\nabla^2 U$ |
|------------------------------|------------------|
| 1) $r^2 (= x^2 + y^2 + z^2)$ | 6 |
| 2) xy^2z^3 | $2xz^3 + 6xy^2z$ |
| 3) $1/r$ | 0 |

Let's prove example (3) (which is particularly significant – can you guess why?).

$$1/r = (x^2 + y^2 + z^2)^{-1/2}$$

$$\begin{aligned} \frac{\partial}{\partial x} \frac{\partial}{\partial x} (x^2 + y^2 + z^2)^{-1/2} &= \frac{\partial}{\partial x} -x \cdot (x^2 + y^2 + z^2)^{-3/2} \\ &= -(x^2 + y^2 + z^2)^{-3/2} + 3x \cdot x \cdot (x^2 + y^2 + z^2)^{-5/2} \\ &= (1/r^3)(-1 + 3x^2/r^2) \end{aligned}$$

Adding up similar terms for y and z

$$\nabla^2 \frac{1}{r} = \frac{1}{r^3} \left(-3 + 3 \frac{(x^2 + y^2 + z^2)}{r^2} \right) = 0$$

5.6 The curl of a vector field

So far we have seen the operator ∇ applied to a scalar field ∇U ; and dotted with a vector field $\nabla \cdot \mathbf{a}$.

We are now overwhelmed by an irresistible temptation to

- cross it with a vector field $\nabla \times \mathbf{a}$

This gives the **curl of a vector field**

$$\nabla \times \mathbf{a} \equiv \text{curl}(\mathbf{a})$$

We can follow the pseudo-determinant recipe for vector products, so that

$$\begin{aligned} \nabla \times \mathbf{a} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix} \quad (\text{remember it this way}) \\ &= \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) \hat{i} + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) \hat{j} + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \hat{k} \end{aligned}$$

♣ Examples of curl evaluation

| \mathbf{a} | $\nabla \times \mathbf{a}$ |
|---|---|
| 1) $-y\hat{\mathbf{i}} + x\hat{\mathbf{j}}$ | $2\hat{\mathbf{k}}$ |
| 2) $x^2y^2\hat{\mathbf{k}}$ | $2x^2y\hat{\mathbf{i}} - 2xy^2\hat{\mathbf{j}}$ |

5.7 The significance of curl

Perhaps the first example gives a clue. The field $\mathbf{a} = -y\hat{\mathbf{i}} + x\hat{\mathbf{j}}$ is sketched in Figure 5.3(a). (It is the field you would calculate as the velocity field of an object rotating with $\boldsymbol{\omega} = [0, 0, 1]$.) This field has a curl of $2\hat{\mathbf{k}}$, which is in the r-h screw sense out of the page. You can also see that a field like this must give a finite value to the line integral around the complete loop $\oint_C \mathbf{a} \cdot d\mathbf{r}$.

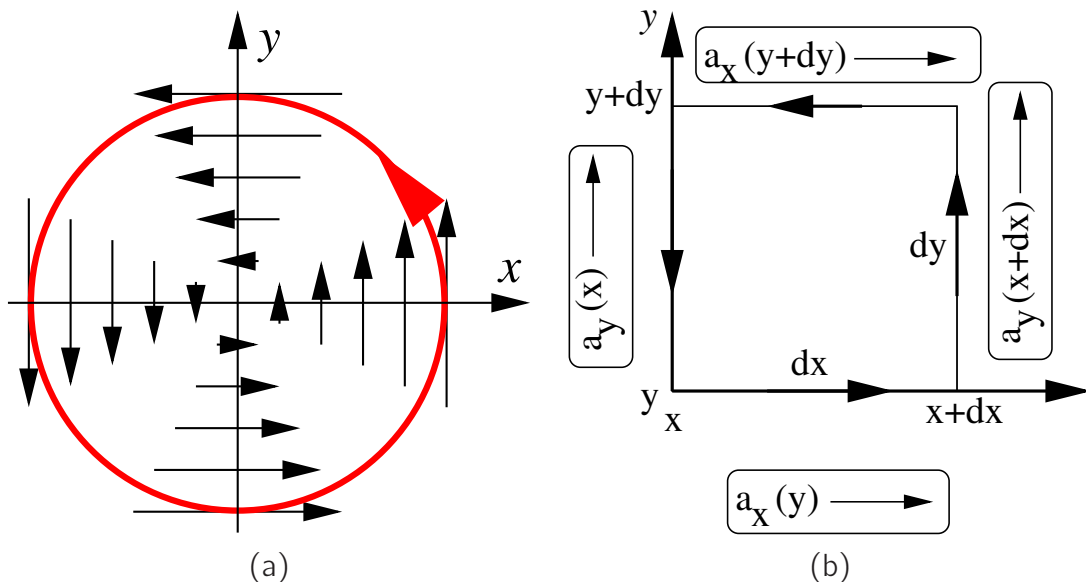


Figure 5.3: (a) A rough sketch of the vector field $-y\hat{\mathbf{i}} + x\hat{\mathbf{j}}$. (b) An element in which to calculate curl.

In fact curl is closely related to the line integral around a loop.

The **circulation** of a vector \mathbf{a} round any closed curve C is defined to be $\oint_C \mathbf{a} \cdot d\mathbf{r}$ and the **curl** of the vector field \mathbf{a} represents the **vorticity**, or **circulation per unit area**, of the field.

Our proof uses the small rectangular element dx by dy shown in Figure 5.3(b). Consider the circulation round the perimeter of a rectangular element.

The fields in the x direction at the bottom and top are

$$a_x(y) \quad \text{and} \quad a_x(y + dy) = a_x(y) + \frac{\partial a_x}{\partial y} dy,$$

where $a_x(y)$ denotes a_x is a function of y , and the fields in the y direction at the left and right are

$$a_y(x) \quad \text{and} \quad a_y(x + dx) = a_y(x) + \frac{\partial a_y}{\partial x} dx$$

Starting at the bottom and working round in the anticlockwise sense, the four contributions to the circulation dC are therefore as follows, where the minus signs take account of the path being opposed to the field:

$$\begin{aligned} dC &= +[a_x(y) dx] + [a_y(x + dx) dy] - [a_x(y + dy) dx] - [a_y(x) dy] \\ &= +[a_x(y) dx] + \left[\left(a_y(x) + \frac{\partial a_y}{\partial x} dx \right) dy \right] - \left[\left(a_x(y) + \frac{\partial a_x}{\partial y} dy \right) dx \right] - [a_y(x) dy] \\ &= \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) dx dy \\ &= (\nabla \times \mathbf{a}) \cdot d\mathbf{S} \end{aligned}$$

where $d\mathbf{S} = dx dy \hat{\mathbf{k}}$.

NB: Again, this is not a completely rigorous proof as we have not shown that the result is independent of the co-ordinate system used.

5.8 Some definitions involving div, curl and grad

- A vector field with zero divergence is said to be **solenoidal**.
- A vector field with zero curl is said to be **irrotational**.
- A scalar field with zero gradient is said to be, er, **constant**.

Lecture 6

Vector Operator Identities

In this lecture we look at more complicated identities involving vector operators. The main thing to appreciate is that the operators behave both as vectors and as differential operators, so that the usual rules of taking the derivative of, say, a product must be observed.

There could be a cottage industry inventing vector identities. HLT contains a lot of them. So why not leave it at that?

First, since grad, div and curl describe key aspects of vector fields, they arise often in practice, and so the identities can save you a lot of time and hacking of partial derivatives, as we will see when we consider Maxwell's equation as an example later.

Secondly, they help to identify other practically important vector operators. So, although this material is a bit dry, the relevance of the identities should become clear later in other Engineering courses.

6.1 Identity 1: curl grad $U = 0$

$$\begin{aligned}\nabla \times \nabla U &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial U/\partial x & \partial U/\partial y & \partial U/\partial z \end{vmatrix} \\ &= \hat{i} \left(\frac{\partial^2 U}{\partial y \partial z} - \frac{\partial^2 U}{\partial z \partial y} \right) + \hat{j} () + \hat{k} () \\ &= \mathbf{0} ,\end{aligned}$$

as $\partial^2/\partial y \partial z = \partial^2/\partial z \partial y$.

Note that the output is a null vector.

6.2 Identity 2: $\text{div curl } \mathbf{a} = 0$

$$\begin{aligned}\nabla \cdot \nabla \times \mathbf{a} &= \begin{vmatrix} \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ a_x & a_y & a_z \end{vmatrix} \\ &= \frac{\partial^2 a_z}{\partial x \partial y} - \frac{\partial^2 a_y}{\partial x \partial z} - \frac{\partial^2 a_z}{\partial y \partial x} + \frac{\partial^2 a_x}{\partial y \partial z} + \frac{\partial^2 a_y}{\partial z \partial x} - \frac{\partial^2 a_x}{\partial z \partial y} \\ &= 0\end{aligned}$$

6.3 Identity 3: div and curl of $U\mathbf{a}$

Suppose that $U(\mathbf{r})$ is a scalar field and that $\mathbf{a}(\mathbf{r})$ is a vector field and we are interested in the product $U\mathbf{a}$. This is a vector field, so we can compute its divergence and curl. For example the density $\rho(\mathbf{r})$ of a fluid is a scalar field, and the instantaneous velocity of the fluid $\mathbf{v}(\mathbf{r})$ is a vector field, and we are probably interested in mass flow rates for which we will be interested in $\rho(\mathbf{r})\mathbf{v}(\mathbf{r})$.

The divergence (a scalar) of the product $U\mathbf{a}$ is given by:

$$\begin{aligned}\nabla \cdot (U\mathbf{a}) &= U(\nabla \cdot \mathbf{a}) + (\nabla U) \cdot \mathbf{a} \\ &= U \text{div} \mathbf{a} + (\text{grad} U) \cdot \mathbf{a}\end{aligned}$$

In a similar way, we can take the curl of the vector field $U\mathbf{a}$, and the result should be a vector field:

$$\nabla \times (U\mathbf{a}) = U\nabla \times \mathbf{a} + (\nabla U) \times \mathbf{a} .$$

6.4 Identity 4: div of $\mathbf{a} \times \mathbf{b}$

Life quickly gets trickier when vector or scalar products are involved: For example, it is not *that* obvious that

$$\text{div}(\mathbf{a} \times \mathbf{b}) = \text{curl} \mathbf{a} \cdot \mathbf{b} - \mathbf{a} \cdot \text{curl} \mathbf{b}$$

To show this, use the determinant:

$$\begin{aligned}\begin{vmatrix} \partial/\partial x_i & \partial/\partial x_j & \partial/\partial x_k \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} &= \frac{\partial}{\partial x} [a_y b_z - a_z b_y] + \frac{\partial}{\partial y} [a_z b_x - a_x b_z] + \frac{\partial}{\partial z} [a_x b_y - a_y b_x] \\ &= \dots \text{bash out the products} \dots \\ &= \text{curl} \mathbf{a} \cdot \mathbf{b} - \mathbf{a} \cdot (\text{curl } \mathbf{b})\end{aligned}$$

6.5 Identity 5: $\text{curl}(\mathbf{a} \times \mathbf{b})$

$$\text{curl}(\mathbf{a} \times \mathbf{b}) = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ a_y b_z - a_z b_y & a_z b_x - a_x b_z & a_x b_y - a_y b_x \end{vmatrix}$$

so the \hat{i} component is

$$\frac{\partial}{\partial y}(a_x b_y - a_y b_x) - \frac{\partial}{\partial z}(a_z b_x - a_x b_z)$$

which can be written as the sum of four terms:

$$a_x \left(\frac{\partial b_y}{\partial y} + \frac{\partial b_z}{\partial z} \right) - b_x \left(\frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) + \left(b_y \frac{\partial}{\partial y} + b_z \frac{\partial}{\partial z} \right) a_x - \left(a_y \frac{\partial}{\partial y} + a_z \frac{\partial}{\partial z} \right) b_x$$

Adding $a_x(\partial b_x/\partial x)$ to the first of these, and subtracting it from the last, and doing the same with $b_x(\partial a_x/\partial x)$ to the other two terms, we find that (you should of course check this):

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = (\nabla \cdot \mathbf{b})\mathbf{a} - (\nabla \cdot \mathbf{a})\mathbf{b} + [\mathbf{b} \cdot \nabla]\mathbf{a} - [\mathbf{a} \cdot \nabla]\mathbf{b}$$

where $[\mathbf{a} \cdot \nabla]$ can be regarded as new, and very useful, scalar differential operator.

6.6 Definition of the operator $[\mathbf{a} \cdot \nabla]$

This is a *scalar operator*, but it can obviously be applied to a scalar field, resulting in a scalar field, or to a vector field resulting in a vector field:

$$[\mathbf{a} \cdot \nabla] \equiv \left[a_x \frac{\partial}{\partial x} + a_y \frac{\partial}{\partial y} + a_z \frac{\partial}{\partial z} \right] .$$

6.7 Identity 6: $\text{curl}(\text{curl} \mathbf{a})$ for you to derive

The following important identity is stated, and left as an exercise:

$$\text{curl}(\text{curl} \mathbf{a}) = \text{grad div} \mathbf{a} - \nabla^2 \mathbf{a}$$

where

$$\nabla^2 \mathbf{a} = \nabla^2 a_x \hat{i} + \nabla^2 a_y \hat{j} + \nabla^2 a_z \hat{k}$$

♣ Example of Identity 6: electromagnetic waves

Q: James Clerk Maxwell established a set of four vector equations which are fundamental to working out how electromagnetic waves propagate. The entire telecommunications industry is built on these.

$$\begin{aligned}\operatorname{div} \mathbf{D} &= \rho \\ \operatorname{div} \mathbf{B} &= 0 \\ \operatorname{curl} \mathbf{E} &= -\frac{\partial}{\partial t} \mathbf{B} \\ \operatorname{curl} \mathbf{H} &= \mathbf{J} + \frac{\partial}{\partial t} \mathbf{D}\end{aligned}$$

In addition, we can assume the following, which should all be familiar to you:
 $\mathbf{B} = \mu_r \mu_0 \mathbf{H}$, $\mathbf{J} = \sigma \mathbf{E}$, $\mathbf{D} = \epsilon_r \epsilon_0 \mathbf{E}$,

where all the scalars are constants.

Now show that in a material with zero free charge density, $\rho = 0$, and with zero conductivity, $\sigma = 0$, the electric field \mathbf{E} must be a solution of the wave equation

$$\nabla^2 \mathbf{E} = \mu_r \mu_0 \epsilon_r \epsilon_0 (\partial^2 \mathbf{E} / \partial t^2) .$$

A: First, a bit of respect. Imagine you are the first to do this — this is a tingle moment.

$$\begin{aligned}\operatorname{div} \mathbf{D} &= \operatorname{div}(\epsilon_r \epsilon_0 \mathbf{E}) = \epsilon_r \epsilon_0 \operatorname{div} \mathbf{E} = \rho = 0 \Rightarrow \operatorname{div} \mathbf{E} = 0. & (a) \\ \operatorname{div} \mathbf{B} &= \operatorname{div}(\mu_r \mu_0 \mathbf{H}) = \mu_r \mu_0 \operatorname{div} \mathbf{H} = 0 \Rightarrow \operatorname{div} \mathbf{B} = 0 & (b) \\ \operatorname{curl} \mathbf{E} &= -\partial \mathbf{B} / \partial t = -\mu_r \mu_0 (\partial \mathbf{H} / \partial t) & (c) \\ \operatorname{curl} \mathbf{H} &= \mathbf{J} + \partial \mathbf{D} / \partial t = \mathbf{0} + \epsilon_r \epsilon_0 (\partial \mathbf{E} / \partial t) & (d)\end{aligned}$$

But we know (or rather you worked out in Identity 6) that $\operatorname{curl} \operatorname{curl} = \operatorname{grad} \operatorname{div} - \nabla^2$, and using (c)

$$\operatorname{curl} \operatorname{curl} \mathbf{E} = \operatorname{grad} \operatorname{div} \mathbf{E} - \nabla^2 \mathbf{E} = \operatorname{curl} (-\mu_r \mu_0 (\partial \mathbf{H} / \partial t))$$

so interchanging the order of partial differentiation, and using (a) $\operatorname{div} \mathbf{E} = 0$:

$$\begin{aligned}-\nabla^2 \mathbf{E} &= -\mu_r \mu_0 \frac{\partial}{\partial t} (\operatorname{curl} \mathbf{H}) \\ &= -\mu_r \mu_0 \frac{\partial}{\partial t} \left(\epsilon_r \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \\ \Rightarrow \nabla^2 \mathbf{E} &= \mu_r \mu_0 \epsilon_r \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}\end{aligned}$$

6.8. GRAD, DIV, CURL AND ∇^2 IN CURVILINEAR CO-ORDINATE SYSTEMS

This equation is actually three equations, one for each component:

$$\nabla^2 E_x = \mu_r \mu_0 \epsilon_r \epsilon_0 \frac{\partial^2 E_x}{\partial t^2}$$

and so on for E_y and E_z .

6.8 Grad, div, curl and ∇^2 in curvilinear co-ordinate systems

It is possible to obtain general expressions for grad, div and curl in any orthogonal curvilinear co-ordinate system by making use of the h factors which were introduced in Lecture 4.

We recall that the unit vector in the direction of increasing u , with v and w being kept constant, is

$$\hat{\mathbf{u}} = \frac{1}{h_u} \frac{\partial \mathbf{r}}{\partial u}$$

where \mathbf{r} is the position vector, and

$$h_u = \left| \frac{\partial \mathbf{r}}{\partial u} \right|$$

is the metric coefficient. Similar expressions apply for the other co-ordinate directions. Then

$$d\mathbf{r} = h_u \hat{\mathbf{u}} du + h_v \hat{\mathbf{v}} dv + h_w \hat{\mathbf{w}} dw .$$

6.9 Grad in curvilinear coordinates

Noting that $U = U(\mathbf{r})$ and $U = U(u, v, w)$, and using the properties of the gradient of a scalar field obtained previously

$$\nabla U \cdot d\mathbf{r} = dU = \frac{\partial U}{\partial u} du + \frac{\partial U}{\partial v} dv + \frac{\partial U}{\partial w} dw$$

It follows that

$$\nabla U \cdot (h_u \hat{\mathbf{u}} du + h_v \hat{\mathbf{v}} dv + h_w \hat{\mathbf{w}} dw) = \frac{\partial U}{\partial u} du + \frac{\partial U}{\partial v} dv + \frac{\partial U}{\partial w} dw$$

The only way this can be satisfied for independent du, dv, dw is when

$$\nabla U = \frac{1}{h_u} \frac{\partial U}{\partial u} \hat{\mathbf{u}} + \frac{1}{h_v} \frac{\partial U}{\partial v} \hat{\mathbf{v}} + \frac{1}{h_w} \frac{\partial U}{\partial w} \hat{\mathbf{w}}$$

6.10 Divergence in curvilinear coordinates

Expressions can be obtained for the divergence of a vector field in orthogonal curvilinear co-ordinates by making use of the flux property.

We consider an element of volume dV . If the curvilinear coordinates are orthogonal then the little volume is a cuboid (to first order in small quantities) and

$$dV = h_u h_v h_w du dv dw .$$

However, it is not quite a cuboid: the area of two opposite faces will differ as the scale parameters are functions of u , v and w in general.

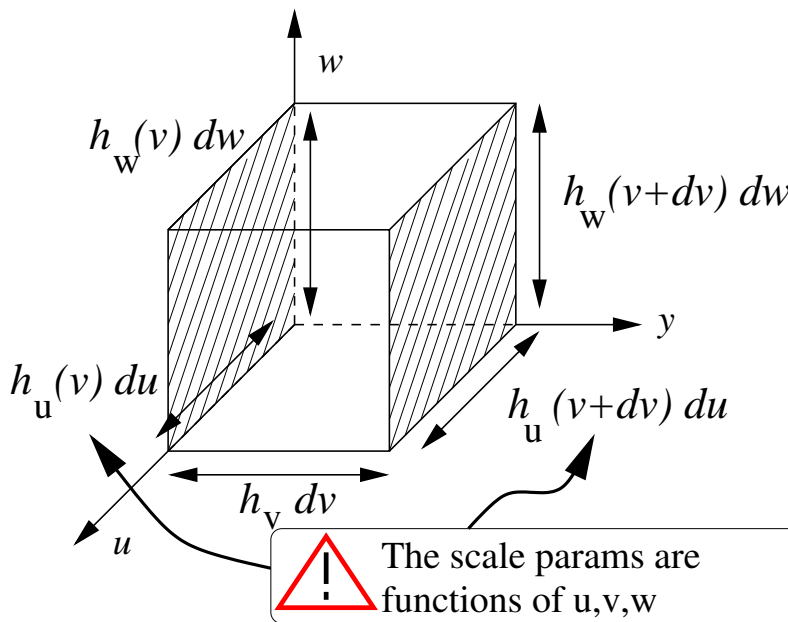


Figure 6.1: Elemental volume for calculating divergence in orthogonal curvilinear coordinates

So the net efflux from the two faces in the $\hat{\mathbf{v}}$ direction shown in Figure 6.1 is

$$\begin{aligned} &= \left[a_v + \frac{\partial a_v}{\partial v} dv \right] \left[h_u + \frac{\partial h_u}{\partial v} dv \right] \left[h_w + \frac{\partial h_w}{\partial v} dv \right] dudw - a_v h_u h_w dudw \\ &= \frac{\partial(a_v h_u h_w)}{\partial v} dudvdw \end{aligned}$$

which is easily shown by multiplying the first line out and dropping second order terms (i.e. $(dv)^2$).

By definition div is the net efflux per unit volume, so summing up the other faces:

$$\begin{aligned} \text{div} \mathbf{a} dV &= \left(\frac{\partial(a_u h_v h_w)}{\partial u} + \frac{\partial(a_v h_u h_w)}{\partial v} + \frac{\partial(a_w h_u h_v)}{\partial w} \right) dudvdw \\ \Rightarrow \text{div} \mathbf{a} h_u h_v h_w dudvdw &= \left(\frac{\partial(a_u h_v h_w)}{\partial u} + \frac{\partial(a_v h_u h_w)}{\partial v} + \frac{\partial(a_w h_u h_v)}{\partial w} \right) dudvdw \end{aligned}$$

6.11. CURL IN CURVILINEAR COORDINATES

So, finally,

$$\text{div} \mathbf{a} = \frac{1}{h_u h_v h_w} \left(\frac{\partial(a_u h_v h_w)}{\partial u} + \frac{\partial(a_v h_u h_w)}{\partial v} + \frac{\partial(a_w h_u h_v)}{\partial w} \right)$$

6.11 Curl in curvilinear coordinates

Recall from Lecture 5 that we computed the z component of curl as the circulation per unit area from

$$dC = \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) dx dy$$

By analogy with our derivation of divergence, you will realize that for an orthogonal curvilinear coordinate system we can write the area as $h_u h_v du dv$. But the opposite sides are no longer quite of the same length. The lower of the pair in Figure 6.2 is length $h_u(v) du$, but the upper is of length $h_u(v + dv) du$

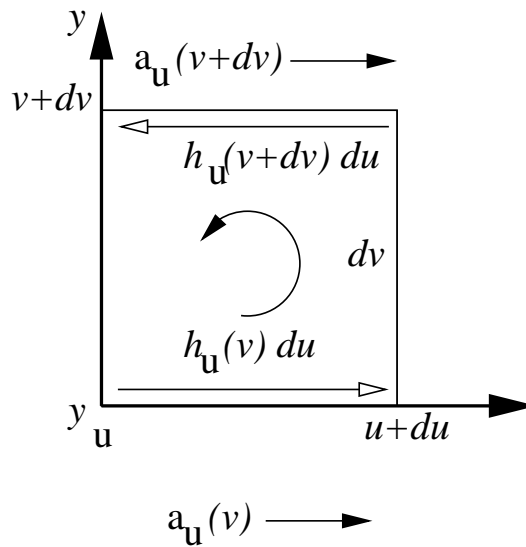


Figure 6.2: Elemental loop for calculating curl in orthogonal curvilinear coordinates

Summing this pair gives a contribution to the circulation

$$a_u(v) h_u(v) du - a_u(v + dv) h_u(v + dv) du = -\frac{\partial(h_u a_u)}{\partial v} dv du$$

and together with the other pair:

$$dC = \left(-\frac{\partial(h_u a_u)}{\partial v} + \frac{\partial(h_v a_v)}{\partial u} \right) du dv$$

So the circulation per unit area is

$$\frac{dC}{h_u h_v du dv} = \frac{1}{h_u h_v} \left(\frac{\partial(h_v a_v)}{\partial u} - \frac{\partial(h_u a_u)}{\partial v} \right)$$

and hence curl is

$$\begin{aligned} \text{curl} \mathbf{a}(u, v, w) = & \frac{1}{h_v h_w} \left(\frac{\partial(h_w a_w)}{\partial v} - \frac{\partial(h_v a_v)}{\partial w} \right) \hat{\mathbf{u}} + \\ & \frac{1}{h_w h_u} \left(\frac{\partial(h_u a_u)}{\partial w} - \frac{\partial(h_w a_w)}{\partial u} \right) \hat{\mathbf{v}} + \\ & \frac{1}{h_u h_v} \left(\frac{\partial(h_v a_v)}{\partial u} - \frac{\partial(h_u a_u)}{\partial v} \right) \hat{\mathbf{w}} \end{aligned}$$

You should check that this can be written as

Curl in curvilinear coords:

$$\text{curl} \mathbf{a}(u, v, w) = \frac{1}{h_u h_v h_w} \begin{vmatrix} h_u \hat{\mathbf{u}} & h_v \hat{\mathbf{v}} & h_w \hat{\mathbf{w}} \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ h_u a_u & h_v a_v & h_w a_w \end{vmatrix}$$

6.12 The Laplacian in curvilinear coordinates

Substitution of the components of $\text{grad} U$ into the expression for $\text{div} \mathbf{a}$ immediately (!*?) gives the following expression for the Laplacian in general orthogonal co-ordinates:

$$\nabla^2 U = \frac{1}{h_u h_v h_w} \left[\frac{\partial}{\partial u} \left(\frac{h_v h_w}{h_u} \frac{\partial U}{\partial u} \right) + \frac{\partial}{\partial v} \left(\frac{h_w h_u}{h_v} \frac{\partial U}{\partial v} \right) + \frac{\partial}{\partial w} \left(\frac{h_u h_v}{h_w} \frac{\partial U}{\partial w} \right) \right].$$

6.13 Grad Div, Curl, ∇^2 in cylindrical polars

Here $(u, v, w) \rightarrow (r, \phi, z)$. The position vector is $\mathbf{r} = r \cos \phi \hat{\mathbf{i}} + r \sin \phi \hat{\mathbf{j}} + z \hat{\mathbf{k}}$, and $h_r = |\partial \mathbf{r} / \partial r|$, etc.

$$\begin{aligned} \Rightarrow h_r &= \sqrt{(\cos^2 \phi + \sin^2 \phi)} = 1, \\ h_\phi &= \sqrt{(r^2 \sin^2 \phi + r^2 \cos^2 \phi)} = r, \\ h_z &= 1 \end{aligned}$$

6.14. GRAD DIV, CURL, ∇^2 IN SPHERICAL POLARS

$$\begin{aligned}\Rightarrow \text{grad}U &= \frac{\partial U}{\partial r} \hat{\mathbf{e}}_r + \frac{1}{r} \frac{\partial U}{\partial \phi} \hat{\mathbf{e}}_\phi + \frac{\partial U}{\partial z} \hat{\mathbf{k}} \\ \text{div} \mathbf{a} &= \frac{1}{r} \left(\frac{\partial(r a_r)}{\partial r} + \frac{\partial a_\phi}{\partial \phi} \right) + \frac{\partial a_z}{\partial z} \\ \text{curl} \mathbf{a} &= \left(\frac{1}{r} \frac{\partial a_z}{\partial \phi} - \frac{\partial a_\phi}{\partial z} \right) \hat{\mathbf{e}}_r + \left(\frac{\partial a_r}{\partial z} - \frac{\partial a_z}{\partial r} \right) \hat{\mathbf{e}}_\phi + \frac{1}{r} \left(\frac{\partial(r a_\phi)}{\partial r} - \frac{\partial a_r}{\partial \phi} \right) \hat{\mathbf{k}} \\ \nabla^2 U &= \text{Tutorial Exercise}\end{aligned}$$

6.14 Grad Div, Curl, ∇^2 in spherical polars

Here $(u, v, w) \rightarrow (r, \theta, \phi)$. The position vector is $\mathbf{r} = r \sin \theta \cos \phi \hat{\mathbf{i}} + r \sin \theta \sin \phi \hat{\mathbf{j}} + r \cos \theta \hat{\mathbf{k}}$.

$$\begin{aligned}\Rightarrow h_r &= \sqrt{(\sin^2 \theta (\cos^2 \phi + \sin^2 \phi) + \cos^2 \theta)} = 1 \\ h_\theta &= \sqrt{(r^2 \cos^2 \theta (\cos^2 \phi + \sin^2 \phi) + r^2 \sin^2 \theta)} = r \\ h_\phi &= \sqrt{(r^2 \sin^2 \theta (\sin^2 \phi + \cos^2 \phi))} = r \sin \theta\end{aligned}$$

$$\begin{aligned}\Rightarrow \text{grad}U &= \frac{\partial U}{\partial r} \hat{\mathbf{e}}_r + \frac{1}{r} \frac{\partial U}{\partial \theta} \hat{\mathbf{e}}_\theta + \frac{1}{r \sin \theta} \frac{\partial U}{\partial \phi} \hat{\mathbf{e}}_\phi \\ \text{div} \mathbf{a} &= \frac{1}{r^2} \frac{\partial(r^2 a_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(a_\theta \sin \theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial a_\phi}{\partial \phi} \\ \text{curl} \mathbf{a} &= \frac{\hat{\mathbf{e}}_r}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (a_\phi \sin \theta) - \frac{\partial}{\partial \phi} (a_\theta) \right) + \frac{\hat{\mathbf{e}}_\theta}{r \sin \theta} \left(\frac{\partial}{\partial \phi} (a_r) - \frac{\partial}{\partial r} (a_\phi r \sin \theta) \right) + \\ &\quad \frac{\hat{\mathbf{e}}_\phi}{r} \left(\frac{\partial}{\partial r} (a_\theta r) - \frac{\partial}{\partial \theta} (a_r) \right) \\ \nabla^2 U &= \text{Tutorial Exercise}\end{aligned}$$

♣ Examples

Q1 Find $\text{curl} \mathbf{a}$ in (i) Cartesians and (ii) Spherical polars when $\mathbf{a} = x(\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}})$.

A1 (i) In Cartesians

$$\text{curl} \mathbf{a} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ x^2 & xy & xz \end{vmatrix} = -z\hat{\mathbf{j}} + y\hat{\mathbf{k}}.$$

(ii) In spherical polars, $x = r \sin \theta \cos \phi$ and $\mathbf{r} = (x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}})$. So

$$\begin{aligned}\mathbf{a} &= r^2 \sin \theta \cos \phi \hat{\mathbf{e}}_r \\ \Rightarrow a_r &= r^2 \sin \theta \cos \phi; \quad a_\theta = 0; \quad a_\phi = 0.\end{aligned}$$

Hence as

$$\begin{aligned}\text{curl } \mathbf{a} &= \frac{\hat{\mathbf{e}}_r}{r \sin \theta} \left(\frac{\partial}{\partial \theta} (a_\phi \sin \theta) - \frac{\partial}{\partial \phi} (a_\theta) \right) + \frac{\hat{\mathbf{e}}_\theta}{r \sin \theta} \left(\frac{\partial}{\partial \phi} (a_r) - \frac{\partial}{\partial r} (a_\phi r \sin \theta) \right) + \frac{\hat{\mathbf{e}}_\phi}{r} \left(\frac{\partial}{\partial r} (a_\theta r) - \frac{\partial}{\partial \theta} (a_r) \right) \\ \text{curl } \mathbf{a} &= \frac{\hat{\mathbf{e}}_\theta}{r \sin \theta} \left(\frac{\partial}{\partial \phi} (r^2 \sin \theta \cos \phi) \right) + \frac{\hat{\mathbf{e}}_\phi}{r} \left(-\frac{\partial}{\partial \theta} (r^2 \sin \theta \cos \phi) \right) \\ &= \frac{\hat{\mathbf{e}}_\theta}{r \sin \theta} (-r^2 \sin \theta \sin \phi) + \frac{\hat{\mathbf{e}}_\phi}{r} (-r^2 \cos \theta \cos \phi) \\ &= \hat{\mathbf{e}}_\theta (-r \sin \phi) + \hat{\mathbf{e}}_\phi (-r \cos \theta \cos \phi)\end{aligned}$$

Checking: these two results should be the same, but to check we need expressions for $\hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_\phi$ in terms of $\hat{\mathbf{i}}$ etc.

Remember that we can work out the unit vectors $\hat{\mathbf{e}}_r$ and so on in terms of $\hat{\mathbf{i}}$ etc using

$$\hat{\mathbf{e}}_r = \frac{1}{h_1} \frac{\partial \mathbf{r}}{\partial r}; \quad \hat{\mathbf{e}}_\theta = \frac{1}{h_2} \frac{\partial \mathbf{r}}{\partial \theta}; \quad \hat{\mathbf{e}}_\phi = \frac{1}{h_3} \frac{\partial \mathbf{r}}{\partial \phi} \quad \text{where } \mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}.$$

Grinding through we find

$$\begin{bmatrix} \hat{\mathbf{e}}_r \\ \hat{\mathbf{e}}_\theta \\ \hat{\mathbf{e}}_\phi \end{bmatrix} = \begin{bmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \\ \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \phi & \cos \phi & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} = [R] \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix}$$

Don't be shocked to see a rotation matrix $[R]$: we are after all rotating one right-handed orthogonal coord system into another.

So the result in spherical polars is

$$\begin{aligned}\text{curl } \mathbf{a} &= (\cos \theta \cos \phi \hat{\mathbf{i}} + \cos \theta \sin \phi \hat{\mathbf{j}} - \sin \theta \hat{\mathbf{k}})(-r \sin \phi) + (-\sin \phi \hat{\mathbf{i}} + \cos \phi \hat{\mathbf{j}})(-r \cos \theta \cos \phi) \\ &= -r \cos \theta \hat{\mathbf{j}} + r \sin \theta \sin \phi \hat{\mathbf{k}} \\ &= -z\hat{\mathbf{j}} + y\hat{\mathbf{k}}\end{aligned}$$

which is exactly the result in Cartesians.

Q2 Find the divergence of the vector field $\mathbf{a} = r\mathbf{c}$ where \mathbf{c} is a constant vector
(i) using Cartesian coordinates and (ii) using Spherical Polar coordinates.

A2 (i) Using Cartesian coords:

$$\begin{aligned}\text{div} \mathbf{a} &= \frac{\partial}{\partial x}(x^2 + y^2 + z^2)^{1/2} c_x + \dots \\ &= x \cdot (x^2 + y^2 + z^2)^{-1/2} c_x + \dots \\ &= \frac{1}{r} \mathbf{r} \cdot \mathbf{c} \quad .\end{aligned}$$

(ii) Using Spherical polars

$$\mathbf{a} = a_r \hat{\mathbf{e}}_r + a_\theta \hat{\mathbf{e}}_\theta + a_\phi \hat{\mathbf{e}}_\phi$$

and our first task is to find a_r and so on. We can't do this by inspection, and finding their values requires more work than you might think! Recall

$$\begin{bmatrix} \hat{\mathbf{e}}_r \\ \hat{\mathbf{e}}_\theta \\ \hat{\mathbf{e}}_\phi \end{bmatrix} = \begin{bmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \\ \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \phi & \cos \phi & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} = [R] \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix}$$

Now the point is the same point in space whatever the coordinate system, so

$$a_r \hat{\mathbf{e}}_r + a_\theta \hat{\mathbf{e}}_\theta + a_\phi \hat{\mathbf{e}}_\phi = a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}}$$

and using the inner product

$$\begin{aligned}\begin{bmatrix} a_r \\ a_\theta \\ a_\phi \end{bmatrix}^\top \begin{bmatrix} \hat{\mathbf{e}}_r \\ \hat{\mathbf{e}}_\theta \\ \hat{\mathbf{e}}_\phi \end{bmatrix} &= \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}^\top \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} \\ \begin{bmatrix} a_r \\ a_\theta \\ a_\phi \end{bmatrix}^\top [R] \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} &= \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}^\top \begin{bmatrix} \hat{\mathbf{i}} \\ \hat{\mathbf{j}} \\ \hat{\mathbf{k}} \end{bmatrix} \\ \Rightarrow \begin{bmatrix} a_r \\ a_\theta \\ a_\phi \end{bmatrix}^\top [R] &= \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}^\top \\ \Rightarrow \begin{bmatrix} a_r \\ a_\theta \\ a_\phi \end{bmatrix}^\top &= \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}^\top [R]^\top \\ \Rightarrow \begin{bmatrix} a_r \\ a_\theta \\ a_\phi \end{bmatrix} &= [R] \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}\end{aligned}$$

For our particular problem, $a_x = rc_x$, etc, where c_x is a constant, so now we can write down

$$\begin{aligned}a_r &= r(\sin \theta \cos \phi c_x + \sin \theta \sin \phi c_y + \cos \theta c_z) \\a_\theta &= r(\cos \theta \cos \phi c_x + \cos \theta \sin \phi c_y - \sin \theta c_z) \\a_\phi &= r(-\sin \phi c_x + \cos \phi c_y)\end{aligned}$$

Now all we need to do is to bash out

$$\text{div} \mathbf{a} = \frac{1}{r^2} \frac{\partial(r^2 a_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(a_\theta \sin \theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial a_\phi}{\partial \phi}$$

In glorious detail this is

$$\begin{aligned}\text{div} \mathbf{a} &= 3(\sin \theta \cos \phi c_x + \sin \theta \sin \phi c_y + \cos \theta c_z) + \\&\quad \frac{1}{\sin \theta} (\cos^2 \theta - \sin^2 \theta)(\cos \phi c_x + \sin \phi c_y) - 2 \sin \theta \cos \theta c_z + \\&\quad \frac{1}{\sin \theta} (-\cos \phi c_x - \sin \phi c_y)\end{aligned}$$

A bit more bashing and you'll find

$$\begin{aligned}\text{div} \mathbf{a} &= \sin \theta \cos \phi c_x + \sin \theta \sin \phi c_y + \cos \theta c_z \\&= \hat{\mathbf{e}}_r \cdot \mathbf{c}\end{aligned}$$

This is EXACTLY what you worked out before of course.

Take home messages from these examples:

- Just as physical vectors are independent of their coordinate systems, so are differential operators.
- Don't forget about the vector geometry you did in the 1st year. Rotation matrices are useful!
- Spherical polars were NOT a good coordinate system in which to think about this problem. Let the symmetry guide you.

Lecture 7

Gauss' and Stokes' Theorems

This section finally begins to deliver on why we introduced div grad and curl. Two theorems, both of them over two hundred years old, are explained:

- **Gauss' Theorem** enables an integral taken over a volume to be replaced by one taken over the surface bounding that volume, and vice versa. Why would we want to do that? Computational efficiency and/or numerical accuracy!
 - **Stokes' Law** enables an integral taken around a closed curve to be replaced by one taken over *any* surface bounded by that curve.
-

7.1 Gauss' Theorem

Suppose that $\mathbf{a}(\mathbf{r})$ is a vector field and we want to compute the total flux of the field across the surface S that bounds a volume V . That is, we are interested in calculating:

$$\int_S \mathbf{a} \cdot d\mathbf{S}$$

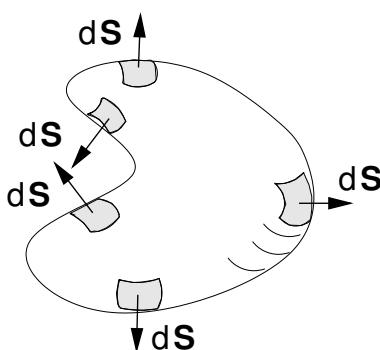


Figure 7.1: The surface element $d\mathbf{S}$ must stick out of the surface.

where recall that $d\mathbf{S}$ is normal to the locally planar surface element and **must everywhere point out of the volume** as shown in Figure 7.1.

Gauss' Theorem tells us that we can do this by considering the total flux generated inside the volume V :

Gauss' Theorem

$$\int_S \mathbf{a} \cdot d\mathbf{S} = \int_V \text{div } \mathbf{a} \, dV$$

obtained by integrating the divergence over the entire volume.

7.2 Informal proof

An non-rigorous proof can be realized by recalling that we defined div by considering the efflux dE from the surfaces of an infinitesimal volume element

$$dE = \mathbf{a} \cdot d\mathbf{S}$$

and defining it as

$$\text{div } \mathbf{a} \, dV = dE = \mathbf{a} \cdot d\mathbf{S}.$$

If we sum over the volume elements, this results in a sum over the surface elements. But if two elemental surface touch, their $d\mathbf{S}$ vectors are in opposing direction and cancel as shown in Figure 7.2. Thus the sum over surface elements gives the overall bounding surface.

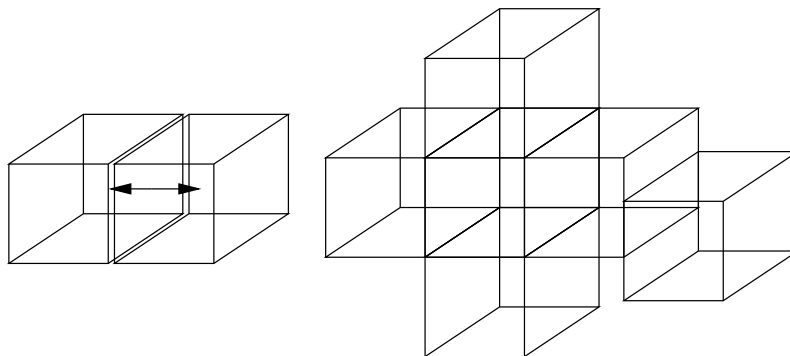


Figure 7.2: When two elements touch, the $d\mathbf{S}$ vectors at the common surface cancel out. One can imagine building the entire volume up from the infinitesimal units.

♣ Example of Gauss' Theorem

This is a typical example, in which the surface integral is rather tedious, whereas the volume integral is straightforward.

Q Derive $\int_S \mathbf{a} \cdot d\mathbf{S}$ where $\mathbf{a} = z^3 \hat{\mathbf{k}}$ and S is the surface of a sphere of radius R centred on the origin:

1. directly;
2. by applying Gauss' Theorem

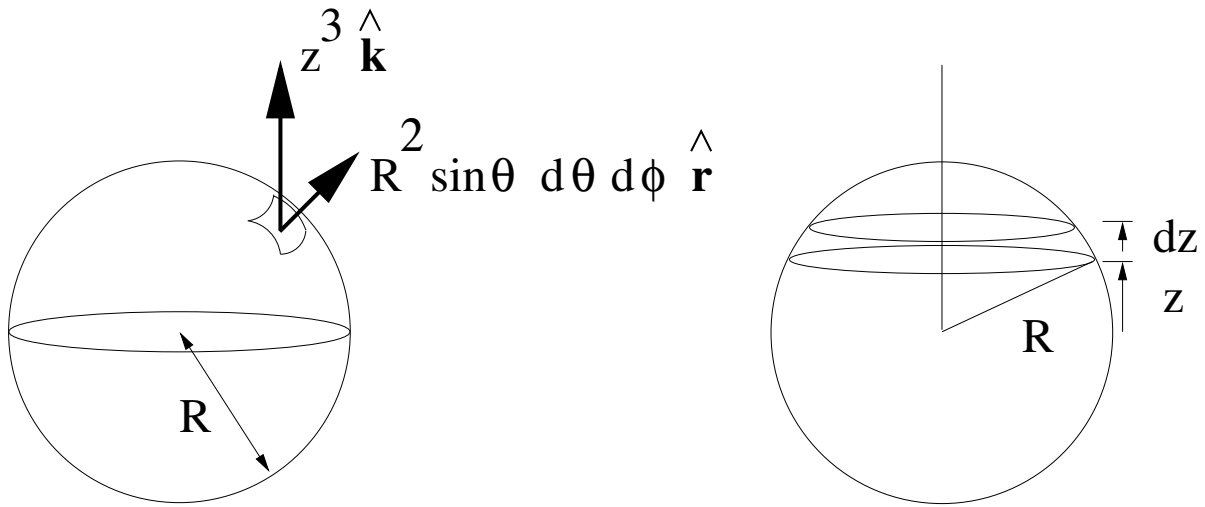


Figure 7.3:

A (1) On the surface of the sphere, $\mathbf{a} = R^3 \cos^3 \theta \hat{\mathbf{k}}$ and $d\mathbf{S} = R^2 \sin \theta d\theta d\phi \hat{\mathbf{r}}$. Everywhere $\hat{\mathbf{r}} \cdot \hat{\mathbf{k}} = \cos \theta$.

$$\begin{aligned}
 \Rightarrow \int_S \mathbf{a} \cdot d\mathbf{S} &= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} R^3 \cos^3 \theta \cdot R^2 \sin \theta d\theta d\phi \hat{\mathbf{e}}_r \cdot \hat{\mathbf{k}} \\
 &= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} R^3 \cos^3 \theta \cdot R^2 \sin \theta d\theta d\phi \cdot \cos \theta \\
 &= 2\pi R^5 \int_0^{\pi} \cos^4 \theta \sin \theta d\theta \\
 &= \frac{2\pi R^5}{5} [-\cos^5 \theta]_0^{\pi} = \frac{4\pi R^5}{5}
 \end{aligned}$$

(2) To apply Gauss' Theorem, we need to figure out $\text{div } \mathbf{a}$ and decide how to compute the volume integral. The first is easy:

$$\text{div } \mathbf{a} = 3z^2$$

For the second, because $\text{div} \mathbf{a}$ involves just z , we can divide the sphere into discs of constant z and thickness dz , as shown in Fig. 7.3. Then

$$dV = \pi(R^2 - z^2)dz$$

and

$$\begin{aligned} \int_V \text{div} \mathbf{a} dV &= 3\pi \int_{-R}^R z^2(R^2 - z^2)dz \\ &= 3\pi \left[\frac{R^2 z^3}{3} - \frac{z^5}{5} \right]_{-R}^R \\ &= \frac{4\pi R^5}{5} \end{aligned}$$

7.3 Surface versus volume integrals

At first sight, it might seem that with a computer performing surface integrals might be better than a volume integral, perhaps because there are, somehow, “fewer elements”. However, this is not the case. Imagine doing a surface integral over a wrinkly surface, say that of the moon. All the elements involved in the integration are “difficult” and must be modelled correctly. With a volume integral, most of the elements are not at the surface, and so the bulk of the integral is done without accurate modelling. The computation is easier, faster, and better conditioned numerically.

7.4 Extension to Gauss' Theorem

Suppose the vector field $\mathbf{a}(\mathbf{r})$ is of the form $\mathbf{a} = U(\mathbf{r})\mathbf{c}$, where $U(\mathbf{r})$ is a scalar field and \mathbf{c} is a constant vector. Then, as we showed in the previous lecture,

$$\begin{aligned} \text{div} \mathbf{a} &= \text{grad} U \cdot \mathbf{c} + U \text{div} \mathbf{c} \\ &= \text{grad} U \cdot \mathbf{c} \end{aligned}$$

since $\text{div} \mathbf{c} = 0$ because \mathbf{c} is constant.

Gauss' Theorem becomes

$$\int_S U \mathbf{c} \cdot d\mathbf{S} = \int_V \text{grad} U \cdot \mathbf{c} dV$$

or, alternatively, taking the constant \mathbf{c} out of the integrals

$$\mathbf{c} \cdot \left(\int_S U d\mathbf{S} \right) = \mathbf{c} \cdot \left(\int_V \text{grad} U dV \right)$$

7.4. EXTENSION TO GAUSS' THEOREM

This is still a scalar equation but we now note that the vector \mathbf{c} is arbitrary so that the result must be true for any vector \mathbf{c} . This can be true only if the vector equation

$$\int_S U d\mathbf{S} = \int_V \text{grad } U dV$$

is satisfied.

If you think this is fishy, just write $\mathbf{c} = \hat{\mathbf{i}}$, then $\mathbf{c} = \hat{\mathbf{j}}$, and $\mathbf{c} = \hat{\mathbf{k}}$ in turn, and you must obtain the three components of $\int_S U d\mathbf{S}$ in turn.

Further “extensions” can be obtained of course. For example one might be able to write the vector field of interest as

$$\mathbf{a}(\mathbf{r}) = \mathbf{b}(\mathbf{r}) \times \mathbf{c}$$

where \mathbf{c} is a constant vector.

♣ Example of extension to Gauss' Theorem

Q $U = x^2 + y^2 + z^2$ is a scalar field, and volume V is the cylinder $x^2 + y^2 \leq a^2$, $0 \leq z \leq h$. Compute the surface integral

$$\int_S U d\mathbf{S}$$

over the surface of the cylinder.

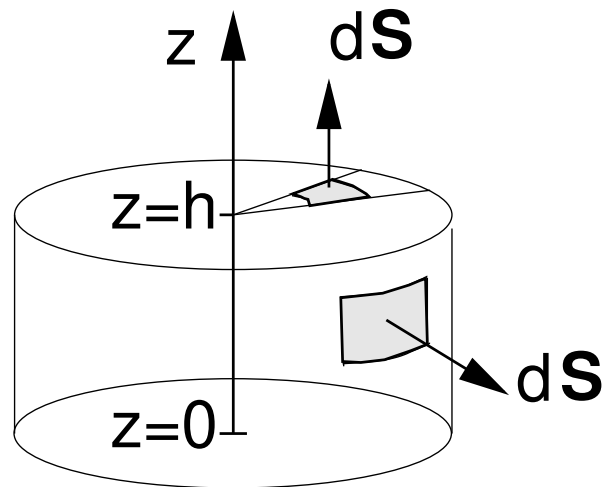
A It is immediately clear from symmetry that there is no contribution from the curved surface of the cylinder since for every vector surface element there exists an equal and opposite element with the same value of U . We therefore need consider only the top and bottom faces.

Top face:

$$U = x^2 + y^2 + z^2 = r^2 + h^2 \text{ and } d\mathbf{S} = r dr d\phi \hat{\mathbf{k}}$$

so

$$\int U d\mathbf{S} = \int_{r=0}^a (h^2 + r^2) 2\pi r dr \int_{\phi=0}^{2\pi} d\phi \hat{\mathbf{k}} = \hat{\mathbf{k}} \pi \left[h^2 r^2 + \frac{1}{2} r^4 \right]_0^a = \pi [h^2 a^2 + \frac{1}{2} a^4] \hat{\mathbf{k}}$$



Bottom face:

$$U = r^2 \text{ and } d\mathbf{S} = -rdrd\phi\hat{\mathbf{k}}$$

The contribution from this face is thus $-\frac{\pi a^4}{2}\hat{\mathbf{k}}$, and the total integral is $\pi h^2 a^2 \hat{\mathbf{k}}$.

On the other hand, using Gauss' Theorem we have to compute

$$\int_V \text{grad } U dV$$

In this case, $\text{grad } U = 2\mathbf{r}$,

$$2 \int_V (x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}) r dr dz d\phi$$

The integrations over x and y are zero by symmetry, so that the only remaining part is

$$2 \int_{z=0}^h z dz \int_{r=0}^a r dr \int_{\phi=0}^{2\pi} d\phi \hat{\mathbf{k}} = \pi a^2 h^2 \hat{\mathbf{k}}$$

7.5 Stokes' Theorem

Stokes' Theorem relates a line integral around a closed path to a surface integral over what is called a *capping surface* of the path.

Stokes' Theorem states:

$$\oint_C \mathbf{a} \cdot d\mathbf{l} = \int_S \text{curl } \mathbf{a} \cdot d\mathbf{S}$$

where S is *any* surface capping the curve C .

Why have we used $d\mathbf{l}$ rather than $d\mathbf{r}$, where \mathbf{r} is the position vector?

There is no good reason for this, as $d\mathbf{l} = d\mathbf{r}$. It just seems to be common usage in line integrals!

7.6 Informal proof

You will recall that in Lecture 5 that we defined curl as the circulation per unit area, and showed that

$$\sum_{\text{around elemental loop}} \mathbf{a} \cdot d\mathbf{l} = dC = (\nabla \times \mathbf{a}) \cdot d\mathbf{S}.$$

Now if we add these little loops together, the internal line sections cancel out because the $d\mathbf{l}$'s are in opposite direction but the field \mathbf{a} is not. This gives the larger surface and the larger bounding contour as shown in Fig. 7.4.

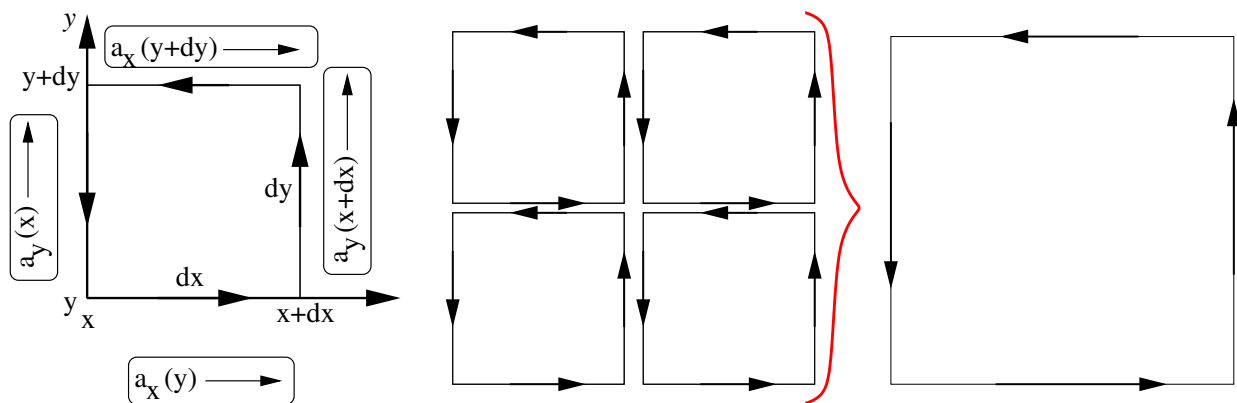


Figure 7.4: An example of an elementary loop, and how they combine together.

For a given contour, the capping surface can be ANY surface bound by the contour. The only requirement is that the surface element vectors point in the “general direction” of a right-handed screw with respect to the sense of the contour integral. See Fig. 7.5.

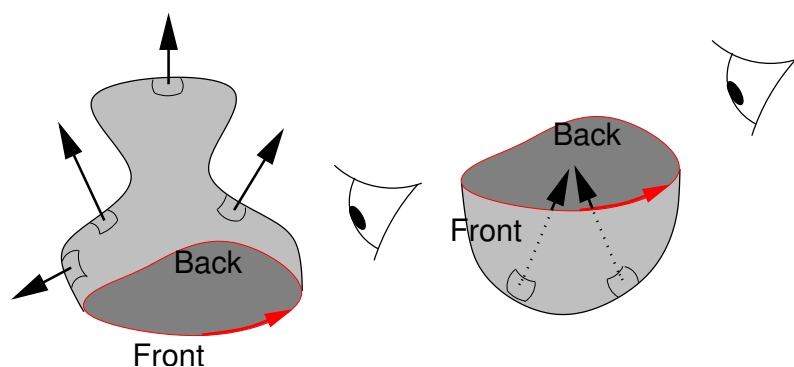


Figure 7.5: For a given contour, the bounding surface can be any shape. $d\mathbf{S}$'s must have a positive component in the sense of a r-h screw wrt the contour sense.

♣ Example of Stokes' Theorem

In practice, (and especially in exam questions!) the bounding contour is often planar, and the capping surface flat or hemispherical or cylindrical.

Q Vector field $\mathbf{a} = x^3\hat{\mathbf{j}} - y^3\hat{\mathbf{i}}$ and C is the circle of radius R centred on the origin. Derive

$$\oint_C \mathbf{a} \cdot d\mathbf{l}$$

directly and (ii) using Stokes' theorem where the surface is the planar surface bounded by the contour.

A(i) Directly. On the circle of radius R

$$\mathbf{a} = R^3(-\sin^3\theta\hat{\mathbf{i}} + \cos^3\theta\hat{\mathbf{j}})$$

and

$$d\mathbf{l} = R d\theta(-\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}})$$

so that:

$$\oint_C \mathbf{a} \cdot d\mathbf{l} = \int_0^{2\pi} R^4(\sin^4\theta + \cos^4\theta) d\theta = \frac{3\pi}{2} R^4,$$

since

$$\int_0^{2\pi} \sin^4\theta d\theta = \int_0^{2\pi} \cos^4\theta d\theta = \frac{3\pi}{4}$$

A(ii) Using Stokes' theorem ...

$$\text{curl } \mathbf{a} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y^3 & x^3 & 0 \end{vmatrix} = 3(x^2 + y^2)\hat{\mathbf{k}} = 3r^2\hat{\mathbf{k}}$$

We choose area elements to be circular strips of radius r thickness dr . Then

$$d\mathbf{S} = 2\pi r dr \hat{\mathbf{k}} \quad \text{and} \quad \int_S \text{curl } \mathbf{a} \cdot d\mathbf{S} = 6\pi \int_0^R r^3 dr = \frac{3\pi}{2} R^4$$

7.7 An Extension to Stokes' Theorem

Just as we considered one extension to Gauss' theorem (not *really* an extension, more of a re-expression), so we will try something similar with Stoke's Theorem.

Again let $\mathbf{a}(\mathbf{r}) = U(\mathbf{r})\mathbf{c}$, where \mathbf{c} is a constant vector. Then

$$\text{curl } \mathbf{a} = U \text{curl } \mathbf{c} + \text{grad } U \times \mathbf{c}$$

Again, $\text{curl } \mathbf{c}$ is zero. Stokes' Theorem becomes in this case:

$$\oint_C U(\mathbf{c} \cdot d\mathbf{l}) = \int_S (\text{grad } U \times \mathbf{c} \cdot d\mathbf{S} = \int_S \mathbf{c} \cdot (d\mathbf{S} \times \text{grad } U)$$

or, rearranging the triple scalar products and taking the constant \mathbf{c} out of the integrals gives

$$\mathbf{c} \cdot \oint_C U d\mathbf{l} = -\mathbf{c} \cdot \int_S \text{grad } U \times d\mathbf{S} .$$

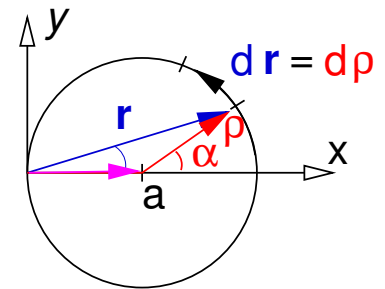
But \mathbf{c} is arbitrary and so

$$\oint_C U d\mathbf{l} = - \int_S \text{grad } U \times d\mathbf{S}$$

7.8 ♣ Example of extension to Stokes' Theorem

Q Derive $\oint_C U d\mathbf{r}$ (i) directly and (ii) using Stokes', where $U = x^2 + y^2 + z^2$ and the line integral is taken around C the circle $(x - a)^2 + y^2 = a^2$ and $z = 0$.

Note that, for no special reason, we have used $d\mathbf{r}$ here not $d\mathbf{l}$.



A(i) First some preamble.

If the circle were centred at the origin, we would write $d\mathbf{r} = ad\theta\hat{\mathbf{e}}_\theta = ad\theta(-\sin\theta\hat{\mathbf{i}} + \cos\theta\hat{\mathbf{j}})$. For such a circle the magnitude $r = |\mathbf{r}| = a$, a constant and so $dr = 0$.

However, in this example $d\mathbf{r}$ is not always in the direction of $\hat{\mathbf{e}}_\theta$, and $dr \neq 0$. Could you write down $d\mathbf{r}$? If not, revise Lecture 3, where we saw that in plane polars $x = r \cos \theta$, $y = r \sin \theta$ and the general expression is

$$d\mathbf{r} = dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} = (\cos\theta dr - r \sin\theta d\theta)\hat{\mathbf{i}} + (\sin\theta dr + r \cos\theta d\theta)\hat{\mathbf{j}}$$

To avoid having to find an expression for r in terms of θ , we will perform a coordinate transformation by writing $\mathbf{r} = [a, 0]^\top + \boldsymbol{\rho}$. So, $x = (a + \rho \cos \alpha)$ and $y = \rho \sin \alpha$, and on the circle itself where $\rho = a$

$$\mathbf{r} = a(1 + \cos \alpha)\hat{\mathbf{i}} + a \sin \alpha \hat{\mathbf{j}} ,$$

$$d\mathbf{r} = a d\alpha (-\sin \alpha \hat{\mathbf{i}} + \cos \alpha \hat{\mathbf{j}}) ,$$

and, as $z = 0$ on the circle,

$$U = a^2(1 + \cos \alpha)^2 + a^2 \sin^2 \alpha = 2a^2(1 + \cos \alpha) .$$

The line integral becomes

$$\oint U d\mathbf{r} = 2a^3 \int_{\alpha=0}^{2\pi} (1 + \cos \alpha)(-\sin \alpha \hat{\mathbf{i}} + \cos \alpha \hat{\mathbf{j}}) d\alpha = 2\pi a^3 \hat{\mathbf{j}}$$

A(ii) Now using Stokes' ...

For a planar surface covering the disc, the surface element can be written using the new parametrization as

$$d\mathbf{S} = \rho d\rho d\alpha \hat{\mathbf{k}}$$

Remember that $U = x^2 + y^2 + z^2 = r^2$, and as $z = 0$ in the plane

$$\text{grad } U = 2(x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}) = 2(a + \rho \cos \alpha)\hat{\mathbf{i}} + 2\rho \sin \alpha \hat{\mathbf{j}} .$$

Be careful to note that x, y are specified for any point on the disc, not on its circular boundary!

So

$$d\mathbf{S} \times \text{grad } U = 2\rho d\rho d\alpha \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 0 & 0 & 1 \\ (a + \rho \cos \alpha) & \rho \sin \alpha & 0 \end{vmatrix} = 2\rho [-\rho \sin \alpha \hat{\mathbf{i}} + (a + \rho \cos \alpha)\hat{\mathbf{j}}] d\rho d\alpha$$

Both $\int_0^{2\pi} \sin \alpha d\alpha = 0$ and $\int_0^{2\pi} \cos \alpha d\alpha = 0$, so we are left with

$$\int_S d\mathbf{S} \times \text{grad } U = \int_{\rho=0}^a \int_{\alpha=0}^{2\pi} 2\rho a \hat{\mathbf{j}} d\rho d\alpha = 2\pi a^3 \hat{\mathbf{j}}$$

Lecture 8

Engineering applications

In Lecture 6 we saw one classic example of the application of vector calculus to Maxwell's equation.

In this lecture we explore a few more examples from fluid mechanics and heat transfer. As with Maxwell's equations, the examples show how vector calculus provides a powerful way of representing underlying physics.

The power comes from the fact that div, grad and curl have a significance or meaning which is more immediate than a collection of partial derivatives. Vector calculus will, with practice, become a convenient shorthand for you.

- Electricity – Ampère's Law
- Fluid Mechanics - The Continuity Equation
- Thermo: The Heat Conduction Equation
- Mechanics/Electrostatics - Conservative fields
- The Inverse Square Law of force
- Gravitational field due to distributed mass
- Gravitational field inside body
- Pressure forces in non-uniform flows

8.1 Electricity – Ampère's Law

If the frequency is low, the displacement current in Maxwell's equation $\text{curl} \mathbf{H} = \mathbf{J} + \partial \mathbf{D} / \partial t$ is negligible, and we find

$$\text{curl} \mathbf{H} = \mathbf{J}$$

Hence

$$\int_S \text{curl} \mathbf{H} \cdot d\mathbf{S} = \int_S \mathbf{J} \cdot d\mathbf{S}$$

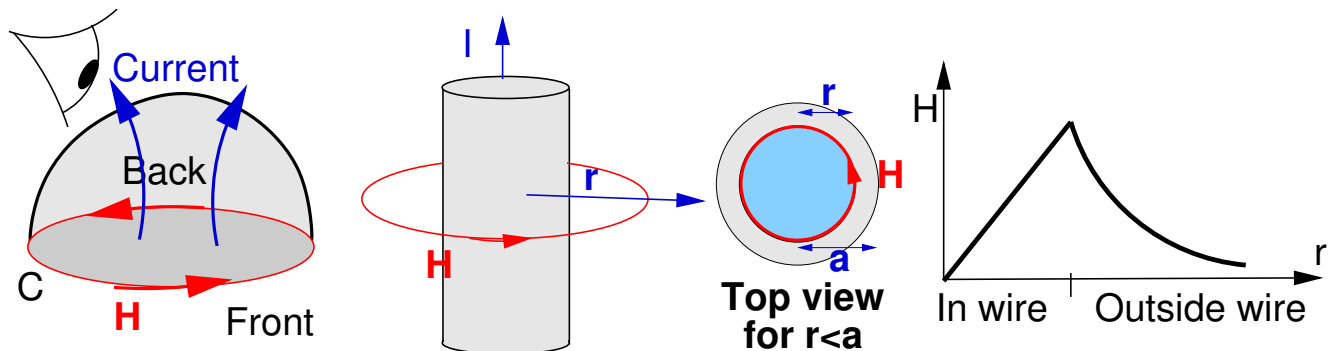
or

$$\oint \mathbf{H} \cdot d\mathbf{l} = \int_S \mathbf{J} \cdot d\mathbf{S}$$

where $\int_S \mathbf{J} \cdot d\mathbf{S}$ is total current through the surface.

Now consider the \mathbf{H} around a straight wire carrying current I . Symmetry tells us the \mathbf{H} is in the $\hat{\mathbf{e}}_\theta$ direction, in a rhs screw sense with respect to the current. (You might check this against Biot-Savart's law.)

Suppose we asked what is the magnitude of \mathbf{H} ?



Inside the wire, the bounding contour only encloses a fraction $(\pi r^2)/(\pi a^2)$ of the current, and so

$$H2\pi r = \int \mathbf{J} \cdot d\mathbf{S} = I(r^2/A^2)$$

$$\Rightarrow H = Ir/2\pi A^2$$

whereas outside we enclose all the current, and so

$$H2\pi r = \int \mathbf{J} \cdot d\mathbf{S} = I$$

$$\Rightarrow H = I/2\pi r$$

A plot is shown in the Figure.

8.2 Fluid Mechanics - The Continuity Equation

The **Continuity Equation** expresses the condition of conservation of mass in a fluid flow. The continuity principle applied to *any* volume (called a *control volume*) may be expressed in words as follows:

“The net rate of mass flow of fluid out of the control volume must equal the rate of decrease of the mass of fluid within the control volume”

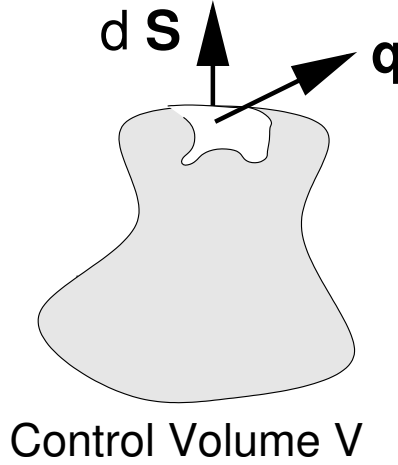


Figure 8.1:

To express the above as a mathematical equation, we denote the velocity of the fluid at each point of the flow by $\mathbf{q}(\mathbf{r})$ (a vector field) and the density by $\rho(\mathbf{r})$ (a scalar field). The element of rate-of-volume-loss through surface $d\mathbf{S}$ is $d\dot{V} = \mathbf{q} \cdot d\mathbf{S}$, so the rate of mass loss is

$$d\dot{M} = \rho \mathbf{q} \cdot d\mathbf{S},$$

so that the total rate of mass loss from the volume is

$$-\frac{\partial}{\partial t} \int_V \rho(\mathbf{r}) dV = \int_S \rho \mathbf{q} \cdot d\mathbf{S}.$$

Assuming that the volume of interest is fixed, this is the same as

$$-\int_V \frac{\partial \rho}{\partial t} dV = \int_S \rho \mathbf{q} \cdot d\mathbf{S}.$$

Now we use Gauss' Theorem to transform the RHS into a volume integral

$$-\int_V \frac{\partial \rho}{\partial t} dV = \int_V \text{div} (\rho \mathbf{q}) dV.$$

The two volume integrals can be equal for any control volume V only if the two integrands are equal at each point of the flow. This leads to the mathematical formulation of

The Continuity Equation:

$$\operatorname{div}(\rho \mathbf{q}) = -\frac{\partial \rho}{\partial t}$$

Notice that if the density doesn't vary with time, $\operatorname{div}(\rho \mathbf{q}) = 0$, and if the density doesn't vary with position then

The Continuity Equation for uniform, time-invariant density:

$$\operatorname{div}(\mathbf{q}) = 0.$$

In this last case, we can say that the flow \mathbf{q} is solenoidal.

8.3 Thermodynamics - The Heat Conduction Equation

Flow of heat is very similar to flow of fluid, and heat flow satisfies a similar continuity equation. The flow is characterized by the heat current density $\mathbf{q}(\mathbf{r})$ (heat flow per unit area and time), sometimes misleadingly called heat flux.

Assuming that there is no mass flow across the boundary of the control volume and no source of heat inside it, the rate of flow of heat out of the control volume by conduction must equal the rate of decrease of internal energy (constant volume) or enthalpy (constant pressure) within it. This leads to the equation

$$\operatorname{div} \mathbf{q} = -\rho c \frac{\partial T}{\partial t},$$

where ρ is the density of the conducting medium, c its specific heat (both are assumed constant) and T is the temperature.

In order to solve for the temperature field another equation is required, linking q to the temperature gradient. This is

$$\mathbf{q} = -\kappa \operatorname{grad} T,$$

where κ is the thermal conductivity of the medium. Combining the two equations gives the *heat conduction equation*:

$$-\operatorname{div} \mathbf{q} = \kappa \operatorname{div} \operatorname{grad} T = \kappa \nabla^2 T = \rho c \frac{\partial T}{\partial t}$$

where it has been assumed that κ is a constant. In steady flow the temperature field satisfies Laplace's Equation $\nabla^2 T = 0$.

8.4 Mechanics - Conservative fields of force

A conservative field of force is one for which the work done

$$\int_A^B \mathbf{F} \cdot d\mathbf{r},$$

moving from A to B is indep. of path taken. As we saw in Lecture 4, conservative fields must satisfy the condition

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = 0,$$

Stokes' tells us that this is

$$\int_S \text{curl } \mathbf{F} \cdot d\mathbf{S} = 0,$$

where S is *any* surface bounded by C .

But if true for *any* C containing A and B, it must be that

$$\text{curl } \mathbf{F} = \mathbf{0}$$

Conservative fields are irrotational
All radial fields are irrotational

One way (actually the only way) of satisfying this condition is for

$$\mathbf{F} = \nabla U$$

The scalar field $U(\mathbf{r})$ is the Potential Function

8.5 The Inverse Square Law of force

Radial forces are found in electrostatics and gravitation — so they are certainly irrotational and conservative.

But in nature these radial forces are also inverse square laws. One reason why this may be so is that it turns out to be the only central force field which is **solenoidal**, i.e. has zero divergence.

If $\mathbf{F} = f(r)\mathbf{r}$,

$$\operatorname{div} \mathbf{F} = 3f(r) + rf'(r).$$

For $\operatorname{div} \mathbf{F} = 0$ we conclude

$$r \frac{df}{dr} + 3f = 0$$

or

$$\frac{df}{f} + 3 \frac{dr}{r} = 0.$$

Integrating with respect to r gives $fr^3 = \text{const} = A$, so that

$$\mathbf{F} = \frac{A\mathbf{r}}{r^3}, \quad |\mathbf{F}| = \frac{A}{r^2}.$$

The condition of zero divergence of the inverse square force field applies everywhere except at $\mathbf{r} = \mathbf{0}$, where the divergence is infinite.

To show this, calculate the outward normal flux out of a sphere of radius R centered on the origin when $\mathbf{F} = F\hat{\mathbf{r}}$. This is

$$\int_S \mathbf{F} \cdot d\mathbf{S} = F \int_{\text{Sphere}} \hat{\mathbf{r}} \cdot d\mathbf{S} = F \int_{\text{Sphere}} d = F4\pi R^2 = 4\pi A = \text{Constant}.$$

Gauss tells us that this flux must be equal to

$$\int_V \operatorname{div} \mathbf{F} dV = \int_0^R \operatorname{div} \mathbf{F} 4\pi r^2 dr$$

where we have done the volume integral as a summation over thin shells of surface area $4\pi r^2$ and thickness dr .

But for all finite r , $\operatorname{div} \mathbf{F} = 0$, so $\operatorname{div} \mathbf{F}$ must be infinite at the origin.

The flux integral is thus

- zero — for any volume which does not contain the origin
- $4\pi A$ for any volume which does contain it.

8.6 Gravitational field due to distributed mass: Poisson's Equation

If one tried the same approach as §8.4 for the gravitational field, $A = Gm$, where m is the mass at the origin and G the universal gravitational constant, one would run into the problem that there is no such thing as point mass.

We can make progress though by considering distributed mass.

The mass contained in each small volume element dV is ρdV and this will make a contribution $-4\pi G \rho dV$ to the flux integral from the control volume. Mass outside the control volume makes no contribution, so that we obtain the equation

$$\int_S \mathbf{F} \cdot d\mathbf{S} = -4\pi G \int_V \rho dV.$$

Transforming the left hand integral by Gauss' Theorem gives

$$\int_V \text{div } \mathbf{F} dV = -4\pi G \int_V \rho dV$$

which, since it is true for any V , implies that

$$-\text{div } \mathbf{F} = 4\pi \rho G.$$

Since the gravitational field is also conservative (i.e. irrotational) it must have an associated potential function U , so that $\mathbf{F} = \text{grad } U$. It follows that the gravitational potential U satisfies

Poisson's Equation

$$\nabla^2 U = 4\pi \rho G.$$

Using the integral form of Poisson's equation, it is possible to calculate the gravitational field inside a spherical body whose density is a function of radius only. We have

$$4\pi R^2 F = 4\pi G \int_0^R 4\pi r^2 \rho dr,$$

where $F = |\mathbf{F}|$, or

$$|F| = \frac{G}{R^2} \int_0^R 4\pi r^2 \rho dr = \frac{MG}{R^2},$$

where M is the total mass inside radius R . For the case of uniform density, this is equal to $M = \frac{4}{3}\pi \rho R^3$ and $|F| = \frac{4}{3}\pi \rho GR$.

8.7 Pressure forces in non-uniform flows

When a body is immersed in a flow it experiences a net pressure force

$$\mathbf{F}_p = - \int_S p d\mathbf{S},$$

where S is the surface of the body. If the pressure p is non-uniform, this integral is not zero. The integral can be transformed using Gauss' Theorem to give the alternative expression

$$\mathbf{F}_p = - \int_V \text{grad } p \, dV,$$

where V is the volume of the body. In the simple hydrostatic case $p + \rho g z = \text{constant}$, so that

$$\text{grad } p = -\rho g \mathbf{k}$$

and the net pressure force is simply

$$\mathbf{F}_p = g \hat{\mathbf{k}} \int_V \rho dV$$

which, in agreement with Archimedes' principle, is equal to the weight of fluid displaced.

